Author:
Edgley, Duncan E

Title:
Ecomorphological diversification during adaptive radiation of East African cichlid fishes

Ecomorphological diversification during adaptive radiation of East African cichlid fishes

DUNCAN E. EDGLEY

School of Biological Sciences

University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy (PhD) in the Faculty of Life Sciences

JANUARY 2023

Supervised by

Professor Martin Genner

Professor Christos Ioannou

Doctor Domino Joyce

Professor George Turner

Main text: 44,560 | Supplementary text: 9,087

Abstract

Adaptive radiations are characterised by the emergence of spectacular morphological diversity over short periods of evolutionary time. A notable example of this is the parallel radiations of haplochromine cichlids in the East African Great Lakes, which have evolved into hundreds of species, in some cases in less than a million years. These cichlids are famous for their diversity in colour, shape and craniofacial morphology, but have also extensively diversified in their sensory systems while adapting to specific niches. However, the evolutionary role of a key sensory modality, the mechanosensory lateral line system, is not known, despite its importance in mediating a range of fish behaviours vital for survival. In this thesis, a combination of comparative morphometric, genomics and behavioural studies are presented, aimed at unpicking the role of lateral line system diversification in cichlid speciation and adaptive radiation. Initially, the remarkable morphological diversity present among the Lake Malawi flock is described. Next, by focusing on ecomorphs within a small crater lake undergoing incipient ecological speciation, it is shown that this disparity is detectable early in population divergence. Following this, the results of research on two hybrid crosses of cichlid species with divergent lateral line system morphologies are presented, providing key insights into the genomic basis of morphological variation, and its significance in mediating collective behaviour. Finally, we quantify natural selection acting on genomic, morphological and ecological traits during undergoing initial population divergence of cichlids. A novel framework for measuring divergent selection on traits in such systems is proposed. Overall, the findings highlight key insights into the significance of this often-overlooked sensory modality for cichlid adaptive radiation. More broadly, these results enhance our understanding of the evolutionary processes that have generated the diversity of ecologically-relevant morphologies that we observe in cichlid radiations.

Cover paintings: Julie Johnson (https://www.lifesciencestudios.com/)

Acknowledgements

First and foremost I'd like to thank my supervisor Martin Genner. It's been a pleasure to work with Martin over the years, and has shown to see how exciting research can be. His help with ideas, data analyses, fieldwork and writing has been invaluable. This PhD was only possible in the first place with his encouragement, and would never have happened without his constant support and guidance.

In Bristol, I'd like to thank the members of our lab group: Zifang Liu, Madeleine Carruthers, Andy Saxon, Harry Layfield and Rupert Collins. Madeleine has been absolutely incredible throughout, with help on fieldwork trips to Tanzania, and with (many) bioinformatics issues in Bristol. Andy is fantastic to work with, and ranting about analysis problems with Harry was incredibly cathartic. I'd also like to thank my various co-supervisors. I'm grateful to Christos Ioannou for his help with behavioural trials, and Domino Joyce for assistance with samples. I'd also like to thank George Turner for his useful discussions, establishing the F_2 cross, and collaborating across many trips to Bangor over the years. I'm also grateful to Elliott Scott and Sabine Hauert for their assistance and collaboration in Chapter 4. I'm incredibly indebted to Emília Santos, whose advice and experience has been incredible over our 9 trips to Bangor since 2018. The PhD would not have been possible without her. I'm also grateful to the members of Emília's lab who have kindly helped over the years, Aaron Hickey, Jake Morris, Bethan Clarke, Aleksandra Marconi and Melany Henot. I thank Richard Durbin, Tyler Linderoth, Bettina Fischer and Moritz Blumer for data contributions and help with analyses. For their help in the aquarium, I'm grateful to Peter Gardiner in Bristol, Alan Smith in Hull, and Michael Hayle in Bangor. So many people have provided data and help during this project, so apologies if there's anyone I've missed. I am grateful to my funder, the NERC GW4+ DTP for making this PhD possible (NE/L002434/1). I'm also grateful to Sara Tonge and Emmiliana Palk for their help with funding when things got tough. A particular thanks has to go to Chris Cammies, a big part of finishing this PhD was the ridiculous amounts of demonstrating.

I'd like to thank every member of my PhD cohort. In no particular order: Jack Greenhalgh, Toby Champneys, Roksana Wilson, Zifang Liu, Rochelle Meah, Dongbo Li, Katie Lihou, Sam England, Luke Romaine, Joanna Attwell, Isla Keesje Davidson, Sverre Tunstad, Pernille Sørensen, Calum Graham, Sarah Dodd, Iestyn Penry-Williams, Hugh Perryman and Anne Kristin-Lenz. I'd also like to thank Molly Clark, Ella Ackroyd, Sophie Carpenter, Duncan O'Brien, Benito Wainwright and Joshua Hufton. Outside of the LSB, I thank Ash, Joe, Jamie, Ollie, Rory, Milo, Fiona, Oscar, Deraj, Ciara, Emily, Emma, Denice, Jag, Elliott, Zoonii and Rebecca. My life for the last 4 years wouldn't have been the same without you.

Lastly, I'd like to thank my family, in particular Mum & Dad for their unwavering support and encouragement over the years. Finally, my biggest thanks go to Matt, for making my life so much better over the final half of this PhD. You kept me going when I didn't think finishing was possible.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

COVID-19 Statement

This thesis is the product of a PhD conducted partially during the lockdowns resulting from COVID-19 pandemic in 2020 and 2021. As a result, the work presented here is significantly different to what was planned during the initial stages of the PhD.

Initially, we hoped that the project would include a broader range of morphological specimens from the Lake Malawi radiation. However, due to the restrictions on international travel, fieldwork was not possible. Therefore, we only included specimens that were obtainable from aquaria at institutions within the UK. In addition, CT scanning and lab work which was vital for Chapter 2, 3 and 4 were delayed due to restrictions on access to wet labs, computer labs and aquaria. Chapter 3 was particularly affected by lockdowns, as it involved regular travel to Bangor University for data collection, which was delayed for a year and a half during the middle of the PhD. The genomic data included in Chapter 3 was delayed significantly, and the full callset was only completed 3 months before the hand-in date. As such, the analyses here were put together very quickly during the writing stage of the thesis. Chapter 5 was borne out of an analysis I conducted during the initial 2020 lockdown while other projects (which were eventually either scrapped or became chapters 2 – 4) were on hold, and this fitness analyses ended up becoming a major theme in my thesis.

Table of Contents

List of Figures

List of Tables

Chapter 1

General Introduction

1.1 Overview

Since the origin of life on Earth, the process of evolution has resulted in a remarkable diversity of forms and morphologies. Central to current evolutionary theory is that much morphological variation is adaptive, in that it is related to survival in the respective environments of different species. Relating the morphology of individual species to its adaptive function has been a longstanding aim of evolutionary biology, and is vital to understanding the dynamics of how populations diversify, isolate and evolve into different species. Nowhere is morphological diversity more evident than during adaptive radiations, in which a multitude of diverse species emerge from common ancestry in short periods of evolutionary time (Gillespie et al. 2020; Schluter 2000). Such adaptive radiations are usually characterised by the emergence of significant morphological diversity, driven by adaptation to different ecological niches. By using adaptive radiations as a model for studying the dynamics of evolutionary change, and mapping behaviour to morphology, and morphology to genome, we can determine how the process of evolution moulds life into the plethora of diverse forms we observe.

1.2 Speciation and Adaptive Radiation

1.2.1 Ecological speciation

Speciation is the evolutionary process by which new species are formed (Coyne and Orr, 2004). Classically this has been defined geographically, with the assumption being that most species are formed when geographic barriers prevent gene flow between populations, leading to accumulation of genomic and phenotypic differences over time (Fitzpatrick et al. 2009; Hernández-Hernández et al. 2021). Although geographical isolation is still thought to play an important role in the diversification process, it is now acknowledged that this is only one of many perspectives from which speciation can be viewed (Coyne and Orr, 2004; Fitzpatrick et al. 2008). For example, it has been found that divergence can often occur with ongoing gene flow in sympatry (Bolnick and Fitzpatrick, 2007; Michel et al. 2010; Malinsky et al. 2015), and along environmental gradients (clines) (Doebeli and Dieckmann, 2003). As well as occurring in an adaptive way through natural selection, it has also been inferred that population differentiation may occur gradually over time through other non-adaptive evolutionary processes, such as genetic drift (Gavrilets 2003; Sobel et al. 2010).

The concept of ecological speciation proposes that adaptation to local environment can lead to disruptive selection on functional traits, eventually leading to reproductive isolation (Rundle and Nosil, 2005). Ecological speciation is thought to be a major driver of morphological diversity during adaptive radiations (Schluter 2000; Rundell and Price, 2009; Stroud and Losos, 2016). Viewing speciation through this lens has led to an increased interest in determining the ecological relevance of morphological disparity, and hence its adaptive function (Wainwright 2007; Coyne and Orr, 2004).

1.2.2 Adaptive radiation

Adaptive radiation is the process by which many species and forms emerge over a relatively short period of evolutionary time. A generally accepted short definition is that adaptive radiation is "the origin of diversity within a rapidly multiplying lineage" (Schluter 2000). Despite there being limited consensus on the specifics of what constitutes an adaptive radiation, they often exhibit consistent characteristics across clades (Schluter 2000; Gillespie et al. 2020). It is generally accepted that there are four features of adaptive radiation: a) common ancestry between species; b) correlation between phenotype and environment; c) functional utility of traits; and d) rapid speciation (Schluter 2000). Most notably, adaptive radiations are almost always accompanied by rapid and extensive diversification in morphology (Schluter 2000).

The classic example of adaptive radiation is Darwin's finches of the Galápagos archipelago (Grant 2003), in which 18 species have evolved beaks adapted to specific trophic niches (Reaney et al. 2020) in just over a million years (Petren et al. 1999; Grant 2003). Other notable examples of macro-scale adaptive radiations include West Indian *Anolis* lizards (Butler et al. 2007), Hawaiian honeycreepers (Tokita et al. 2017), Antarctic icefishes (Wilson et al. 2013), the haplochromine cichlids of the Great Lakes of East Africa (Malinsky et al. 2018; Ronco et al. 2021) and the Neotropical cichlids of Central/South America (Kautt et al. 2020). Other smaller radiations are often referred to as adaptive radiations, including three-spined sticklebacks (Magalhaes et al. 2016), and smaller cichlid species clusters, such as those in soda lakes (Ford et al., 2016) or crater lakes (Barluenga et al. 2006; Martin 2012; Moser et al. 2018).

The circumstances required for adaptive radiation to occur have long been debated, and are still not fully resolved (Gillespie et al. 2020). Some lineages appear to be more prone to radiating, such as haplochromine cichlids and three-spined sticklebacks (Schluter 1993; Seehausen 2006) whereas

others do not – including other cichlid and stickleback lineages (Seehausen 2007; Sparks 2008). This begs the question of whether there are factors intrinsic to certain species or clades that predispose some taxa to radiation, or whether it is entirely a matter of chance and opportunity (Gillespie et al. 2020). One factor suggested to influence the propensity to radiate is the degree of polymorphism within subspecies/subpopulations. For example, if a trait needed to exploit a new niche already exists within a lineage, or is achievable through phenotypic plasticity, this lineage can act to rapidly exploit ecological opportunity (Schluter 2000). One such example of this is in Darwin's finches, in which it has been proposed that the variable beak morphology of mainland ancestors was important for seeding the Galápagos adaptive radiation, by providing a basis on which species could readily adapt to trophic niches (Reaney et al. 2020).

Related to this, it has been proposed that the presence of standing genetic variation underlying a trait can influence propensity to radiate (Schneider and Meyer, 2017; Lai et al. 2019). In a way analogous to existing morphology enabling rapid diversification, if the genomic basis for a phenotype under selection already exists, this can facilitate the process of niche adaptation (Burke 2012). Furthermore, the concept of "key innovations" has been proposed, where specific novel morphology or behaviour phenotypes promote the radiation of species (Hunter 1998; Dumont et al. 2012). It is argued that repeatedly radiating lineages, such as haplochromine cichlid fishes, may have certain traits which facilitate rapid evolutionary change and adaptation (Salzburger et al. 2005).

However, even in radiation-prone lineages, radiation is not guaranteed (Seehausen 2015) and there are likely to be other factors that are critical for rapid clade diversification (Seehausen 2006; Wagner et al. 2012). It has long been thought that the concept of "ecological opportunity" is important for the proliferation of species, in providing a range of niches for species to adapt into (Simpson 1953; Schluter 2000; Stroud and Losos, 2016). This is supported by evidence that adaptive radiations occur commonly in lakes and islands/archipelagos, or following extinction events, all of which have unexploited niche space (Schluter 2000; Stroud and Losos, 2016). Environmental heterogeneity is also thought to be a predictor of biodiversity, in influencing whether circumstances are right for ecological diversification to occur (Rainey and Travisano 1998). As such, ecologically-mediated selection (Schluter and Nagel, 1995; Schluter et al. 1996) driving ecological speciation is thought to be a mechanism by which much functional diversity arises during adaptive radiation (Rundle and Nosil, 2005). In addition, the morphological disparity that

arises during the ecological speciation process is often in ecologically-relevant traits (Schluter 2000).

Rates of cladogenesis vary substantially both between lineages (Seehausen 2015; Takayama et al. 2018) and through time (Gould and Eldredge, 1977), with lineage accumulation generally taking place rapidly when opportunity arises, but otherwise occurring slowly. More abstractly, adaptive radiations can be considered as the inverse of mass-extinction events (Stroud and Losos, 2016). Such rapid loss of biodiversity results in a plethora of ecological opportunities for surviving lineages, and as such, species best placed to do so can radiate, rapidly diversifying to adapt and fill available niches as they emerge. In this way, adaptive radiations may not be rare or unusual events, but in fact may be responsible for much of the diversity we observe.

1.2.3 Measuring natural selection on morphological traits

A longstanding problem in evolutionary biology has been how to quantify natural selection (Pearson 1902; Lande and Arnold, 1983). It is thought to be the primary process by which adaptation occurs during evolution, since proposed independently by Darwin and Russell-Wallace in the mid 19th century (Darwin and Wallace, 1858). As such, key to understanding the dynamics of evolutionary change is measuring and comparing strengths of selection both between clades and between traits (Lande and Arnold, 1983; Hoekstra et al. 2001; Kingsolver et al. 2001).

In 1983, Lande and Arnold formalised quantitative methods to compare the relative strengths of relationships between traits and fitness: the selection gradient and the selection differential (Lande and Arnold, 1983). Selection gradients describe the relationship between a trait and the individual's relative fitness through regression and can be linear (directional selection) (Figure 1.1A), quadratic (disruptive or stabilising selection) (Figure 1.1B & 1.1C), or even have higher order relationships (Moorad and Wade, 2013). Selection differentials describe the difference between the mean trait value before selection and after selection (Lande and Arnold, 1983; Brodie et al. 1995; Kingsolver et al. 2001). These methods form consistent and quantifiable patterns when plotting phenotype against relative fitness, in what is known as a fitness surface (Schluter and Nychka, 1994; Lande and Arnold, 1983) (Figure 1.1A-C).

Lande and Arnold's approaches have led to an expansion of empirical studies in a range of systems measuring selection in the wild, and now selection gradients and differentials are the conventional methods for mapping fitness to phenotype (Hoekstra et al. 2001; Kingsolver et al. 2001; Benkman 2003; Martin and Wainwright, 2013). When measuring selection in wild populations, these

methods have been successful at detecting both linear and non-linear selection, but the magnitude of this selection in both cases is generally quite weak (Hoekstra et al. 2001; Kingsolver et al. 2001). Even though these methods are robust in a multitude of systems, they can be biased by choice of fitness proxy (Viblanc et al. 2022; Franklin and Morrissey, 2017), measurement error, and/or sample size limitations (Kingsolver et al. 2001; Travis 1989).

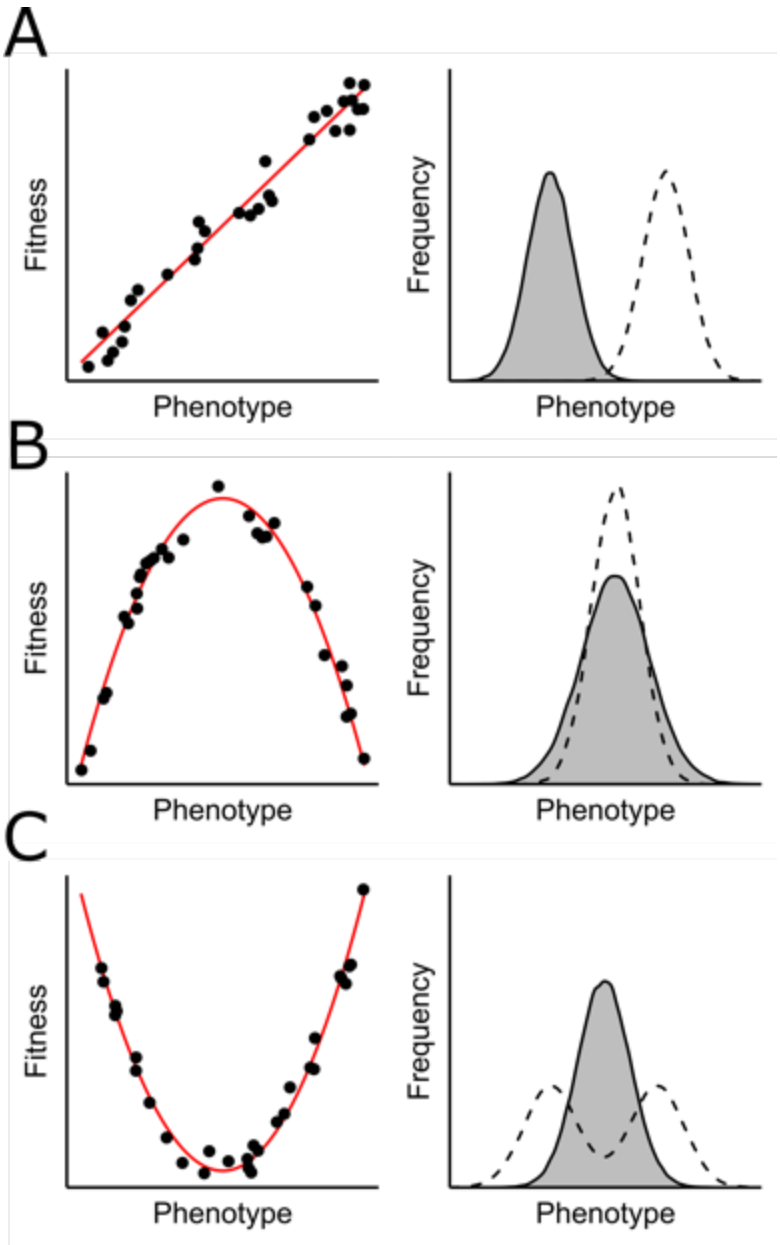


Figure 1.1 Measuring natural selection using selection gradients (Lande and Arnold, 1983). Left are hypothetical fitness surfaces for three evolutionary scenarios, and right are histograms of hypothetical changes in the frequencies of phenotypes in populations under these respective selection regimes: (A) positive directional selection; (B) stabilising selection; and (C) disruptive selection. Grey is the population before selection, and dashed lines represent populations after selection.

As selection affects many traits, phenotype-fitness relationships are often extrapolated to three dimensions to form "fitness landscapes" or "adaptive landscapes", which can be used to visualise the complexity of selection operating at the population level, across the whole genome (Patton et al. 2022), or in gross phenotypic space (Martin and Wainwright, 2013). Adaptive landscapes are also widely used in theoretical models of speciation (de Visser and Krug, 2014; Obolski et al. 2018). During adaptive radiation, we may expect both genomic and morphological fitness landscapes to be rugged and complex, with multiple fitness peaks and troughs, reflecting disruptive selection on genes and morphological traits (Kauffman and Levin, 1987; Schluter 2000; Martin and Wainwright, 2013; Patton et al. 2022). Empirical studies have found that multiple peaks on adaptive landscapes are indicative of divergent evolution, perhaps leading to speciation. As such, they may be used as indicators of ongoing diversification when studying adaptive radiations in the wild (Martin and Wainwright, 2013; Martin and Gould, 2020).

1.2.4 Adaptive radiation of the cichlid fishes of the East African Great Lakes

Cichlids are thought to be one of the most diverse and speciose families of vertebrates, with an estimated 2000 to 3000 species globally (Seehausen 2006; Ronco et al. 2020). To evolutionary biologists, however, it is not only their species richness, but their propensity for rapid bursts of speciation that makes them of such interest. The haplochromine cichlids of the East African Rift Valley lakes are among the most famous examples of adaptive radiations, almost unparalleled in their magnitude (Seehausen 2006). Large-scale African radiations range in age from the mega-radiation within Lake Victoria, in which ~500–600 species have emerged within less than 500,000 years (Seehausen 2006; Meier et al. 2018), to Lake Tanganyika, in which ~200 species have emerged in 5-10 million years (Ronco et al. 2021). The most speciose of these radiations is the Lake Malawi flock, in which ~600–1000 species have evolved within the past one million years (Genner et al. 2007; Malinsky et al. 2018; Svardal et al. 2020) (Figure 1.2).

These East African lacustrine radiations have been extensively studied and are now model systems used by evolutionary biologists to investigate the dynamics of how populations diversify (Seehausen 2006). The presence of multiple radiations occurring in parallel also provides the opportunity to study radiation dynamics and the role of convergent evolution. There is evidence to suggest, for example, that cichlid radiations follow a consistent and predictable pattern (Streelman and Danley, 2003; Muschick et al. 2014; Ronco et al. 2021). It is particularly notable that there is a remarkable consistency in the range of phenotypes that have emerged independently

across Lake Victoria, Lake Malawi and Lake Tanganyika, in body shape, feeding morphology, and pigmentation (Albertson and Kocher, 2006; Muschick et al. 2011; Seehausen 2015).

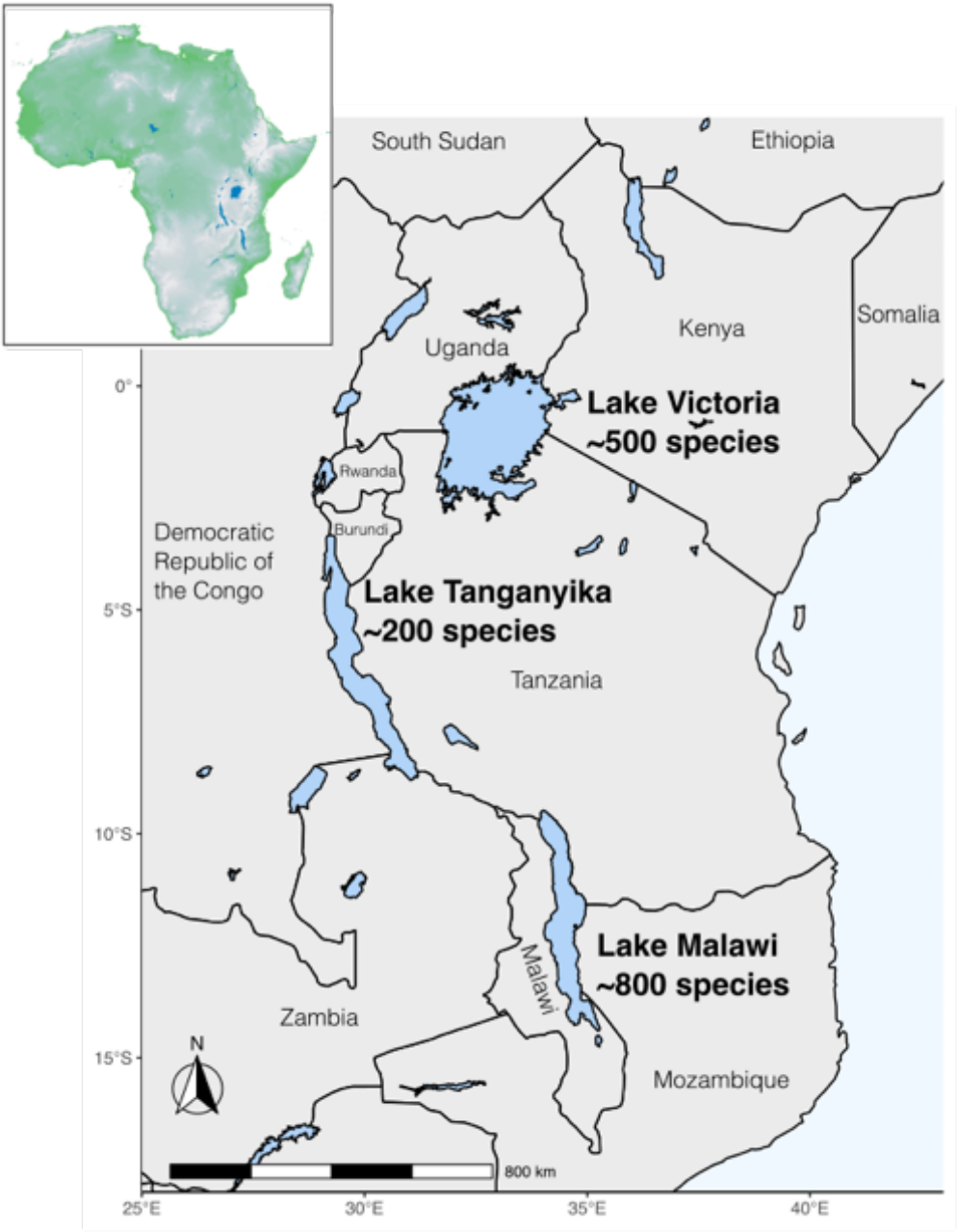


Figure 1.2 Adaptive radiations of the cichlid fishes of the East African Great Lakes. Highlighted are the three Great Rift Valley Lakes of East Africa in which major radiations have occurred, alongside estimates of species richness: Lake Victoria (10-50kya); Lake Malawi (~1mya); and Lake Tanganyika (7-10mya).

1.2 Morphological diversification during adaptive radiation

One of the most striking features of adaptive radiations is the rapid emergence of morphological diversity (Schluter 2000), which tends to be primarily within ecologically-linked functional traits, such as those involved with feeding (Parsons et al. 2011), or sensory perception (Terai et al. 2017;

Edgley and Genner, 2019), but can also involve traits under sexual selection (Salzburger et al. 2005; Selz et al. 2016). Such morphological diversification very often occurs in parallel with population divergence during speciation events (Liu et al. 2018; Struck et al. 2018).

1.3.1 Morphological "key innovations"

It has been proposed that "key innovations" may be important in facilitating adaptive radiations in the wild. These are phenotypes have been suggested to promote diversification and speciation, by allowing species to efficiently exploit novel niches (Hunter 1998; Schluter 2000; Seehausen 2006; Miller et al. 2022). Importantly, such key innovations in morphology are not necessarily morphological novelties (Erwin 2015) but may simply take the form of plasticity in phenotype (Nonaka et al. 2015) or the absence/removal of certain evolutionary constraints (Sharov 2014; Furusawa and Irie, 2020). In African haplochromine cichlids, potential key innovations in morphology include the decoupling of the oral and pharyngeal jaws (Wainwright 2007; Conith and Albertson, 2021), male eggspots (Salzburger 2009; Santos et al. 2014), mouthbrooding and associated morphologies (Duponchelle et al. 2008) and variation in male nuptial colouration (Salzburger et al. 2005; Maan and Sefc, 2013; Kratochwil et al. 2018).

1.3.2 Morphological diversification during cichlid adaptive radiation

Cichlid adaptive radiations are characterised by a spectacular diversity in morphology (Seehausen 2006). In many key cichlid traits, morphological disparity is ecologically-linked, meaning the observed diversity in functional traits is related to pressures from natural selection (Takahashi and Koblmüller, 2011). For instance, in parallel across all three East African Great Lakes, cichlids have evolved a variety of body shapes and sizes (Navon et al. 2017; DeLorenzo et al. 2021), generally following the pelagic-benthic axis of longer and more slender bodies in the open water (e.g. *Rhamphochromis* in Lake Malawi (Turner et al. 2002; Maeda et al. 2009)), and more laterally compressed, deeper bodies in the benthic zone, as is commonly observed in all fish clades (Friedman et al. 2020; DeLorenzo et al. 2021).

Another major axis of cichlid fish diversification is craniofacial diversity and the structure of the cichlid jaw. The morphology of craniofacial structures is tightly linked to trophic niche, and is thought to be primarily driven by adaptation to different diets (Hulsey and García de León, 2005). Adaptation of the feeding anatomy of the head has also been found to show consistency when transitioning between the benthic and pelagic zones across multiple cichlid radiations, reflecting the different trophic niches present along key ecological gradients (Cooper et al. 2010). In addition

to broad dietary specialisation into piscivores, zooplanktivores, algivores and substrate feeders (Edgley and Genner, 2019), other more specific morphologies have emerged convergently across radiations. For example, some algivorous cichlids have evolved flat and wide jaws for scraping algae from rocks (Genner and Turner, 2005; Ford et al. 2016; Edgley and Genner, 2019). Others have evolved large lips (e.g. *Haplochromis chilotes* of the Lake Victoria radiation) as a novel foraging adaptation for finding food in rocky crevices (Baumgarten et al. 2015). In addition, convergently in both Lake Malawi and Lake Tanganyika, cichlids have evolved specialised sensory morphologies in their cranial lateral line system for feeding on buried prey items, within the *Aulonocara* and *Trematocara* genera respectively (Takahashi 2002; Schwalbe et al. 2012). Convergent evolution of adaptive phenotypes is common across parallel East African cichlid radiations (Seehausen 2006), but some axes of diversification may only occur in certain radiations, depending on available niche space. For example, in Lake Malawi, some cichlids have evolved a pelagic planktivorous or piscivorous habit (e.g. *Diplotaxodon* and *Rhamphochromis*) (Thompson et al. 1996). In Lake Tanganyika, however, despite the emergence of some deep water pelagic predatory cichlid species, the analogous niches to Lake Malawi's midwater *Rhamphochromis* species are primarily filled by Clupeidae and Latidae (Kmentová et al. 2020).

Being such an important trait for adaptation to trophic niches, the jaw morphologies of cichlids have been well characterised, using both morphometric and comparative genomic analyses (Albertson et al. 2003; Hulsey and García de León, 2005). Cichlids have decoupled oral and pharyngeal jaws, with the oral jaws being used for prey capture and the pharyngeal jaws for prey processing (Muschick et al. 2011; Ronco and Salzburger, 2021). It is thought that the decoupling of the structures may have enabled the two structures to operate as distinct modules, thus removing the evolutionary constraint imposed by the morphological linkage between the two structures. This may also have allowed them to follow independent trajectories of morphological change over time, and perhaps facilitated a rapid rate of speciation. Despite this, recent studies have found that the two systems are genomically coupled (Conith and Albertson, 2021), which may mean the apparent independent evolutionary trajectories of the two structures are rather more influenced by plasticity in phenotype (Muschick et al. 2011).

Cichlid lower pharyngeal jaws in particular show significant structural variation across clades. Some cichlids have robust, wide pharyngeal jaws with molariform teeth specialised for crushing arthropods or molluscs. Other species that are primarily herbivorous or piscivorous have pharyngeal jaws that are more delicate, with thinner bones and smaller teeth (Burress 2016).

Pharyngeal jaws have diversified along this trophic axis independently in both the Tanganyika and Malawi radiations (Muschick et al. 2011; Edgley and Genner, 2019), but also in Neotropical cichlid lineages (Hulsey et al. 2020). In addition, these morphological differences are readily detectable during the speciation process, between populations with relatively marginal dietary differences (Malinsky et al. 2015; Vernaz et al. 2022). In terms of other aspects of feeding morphology, cichlid gills also vary substantially, and their structure is closely linked to both habitat and feeding ecology (Galis and Barel, 1979).

In cichlids, perhaps the most apparent traits to have diversified among species are pigmentation and patterning, leading to a multitude of colour forms in all three African Great Lakes (Seehausen et al. 1999; Salzburger 2009; Miyagi et al. 2012; Maan and Sefc, 2013; Terai et al. 2017). Colouration in haplochromine cichlids is brightest and most variable in males, with a large body of evidence suggesting that these colours are sexually selected traits (Selz et al. 2016). Males use colour information in male-male competition (Seehausen and Schluter, 2004), and female cichlids have strong colouration-related mating preferences (Selz et al. 2014; Malinsky et al. 2015). As such, male nuptial colouration has been proposed as a mechanism by which reproductive isolation can be established even with relatively small genetic changes, such as mutations in pigmentation genes, rapidly reducing the timescale in which speciation can occur (Maan et al. 2004; Seehausen et al. 2008; Maan et al. 2017). Achieving reproductive isolation so rapidly in this way is one potential explanation for why haplochromine cichlid fishes have radiated extensively (Seehausen et al. 2008; Carleton et al. 2005).

Cichlid sensory systems are important aspects of adaptation to different ecologies (Plenderleith et al. 2005; Maan et al. 2017; Edgley and Genner, 2019). The most extensively studied sensory system is vision, due to its importance in mate choice, and the wide range of male breeding colours typically present in cichlid fish radiations (Maan and Sefc, 2013; Seehausen et al. 1999). Cichlids vary significantly in their opsin gene coding sequences and expression levels, and hence visual pigments and visual capabilities (Carleton et al. 2005; Parry et al. 2005). Evidence suggests there are two major selection pressures driving adaptation of visual systems. The first is the visual environment, specifically background light level at different depths, and how this relates to key ecological tasks such as prey acquisition and predator avoidance (Sabbah et al. 2010; Malinsky et al. 2015). The second major selection pressure is sexual selection, specifically related to visual perception of male colouration (Maan et al. 2004; Carleton et al. 2005; Selz et al. 2014). Regarding the latter, there is evidence that differentiation of the visual system in cichlid fishes can promote

premating reproductive isolation between populations through sensory drive (Schluter and Price, 1993; Carleton et al. 2005; Seehausen et al. 2008). Another key sensory system in fishes, the mechanosensory lateral line, is less well studied in cichlids, with most research focusing on a few individual species with highly specific morphologies (Webb et al. 2014). In Lake Tanganyika and Lake Victoria, the extent of variation in the system is not known, and recent evidence suggests a much more extensive morphological variation across the Lake Malawi radiation than had been previously assumed (Edgley and Genner, 2019).

It is apparent that morphological evolution during cichlid radiations can be ecologically or sexually driven, with both likely acting simultaneously (Gante and Salzburger, 2012; Wagner et al. 2012) The exact interplay between the two is not fully understood (Seehausen 2006), although some stage-based models for adaptive radiation suggest that initial change is related to broad habitat specialisation, followed by trophic niche adaptation, and then finally within-niche sexual selection that drives divergence of ecologically similar species with divergent coloration (Streelman and Danley 2003). Though there is some evidence that this may be the case broadly, its consistency across parallel radiations is still under debate (Danley and Kocher, 2001; Muschick et al. 2014; Ronco et al. 2021).

1.3.3 The genomic basis of morphological diversity in cichlids

Over the past ~20-30 years, genomics has revolutionised the field of evolutionary biology (Seehausen et al. 2014). In speciation and adaptive radiation research, genome-based approaches have been developed to determine the genetic basis for morphological differences (Malinsky et al. 2018; De-Kayne et al. 2022). Genome-based methods are also regularly used to reconstruct the dynamics and tempo of the radiation process, from single speciation events (Barluenga et al. 2006; Malinsky et al. 2015) to whole adaptive radiations (Malinsky et al., 2018; McGee et al. 2020; Ronco et al. 2021).

Discovery of functional genetic variation involves the mapping of genotype to phenotype, through genome-wide association studies (GWAS) of wild populations (Munby et al. 2021; Yoshida and Yáñez, 2021), quantitative trait locus (QTL) mapping of hybrid crosses (Albertson et al. 2003; Marques et al. 2018) or gene expression analyses (Carruthers et al. 2022; Vernaz et al. 2022). In cichlids, these methods have given insights into the genetic basis of morphological variation in cichlid jaws and craniofacial morphology (Albertson and Kocher, 2006; Navon et al. 2017). In addition, researchers have identified the genomic basis for colour and patterning polymorphism

in several cichlid lineages (Hofmann et al. 2009; O'Quin et al. 2013), confirmed by CRISPR/Cas9 gene editing (Kratochwil et al. 2018; Clark et al. 2022). A functional role for genes of interest can also be confirmed by quantifying expression profiles of focal tissues (El Taher et al. 2021; Carruthers et al. 2022), with precise locations of expression able to be determined using in-situ hybridisation.

1.3.4 Avenues for future research

More research is needed on the adaptive function of morphological variation in cichlids (Takahashi and Koblmüller, 2011), as there are multiple aspects of morphology that tend to be poorly studied. Of all sensory systems, vision is by far the most well-studied, in part due to its relevance for mate-choice and potential for driving diversification (Seehausen et al. 2008). The morphology of the lateral line system, for example, has only recently begun to be investigated in a radiation-wide context (Edgley and Genner, 2019), and the genomic basis of the diversity is unknown. Olfactory cues have also been found to inform mate-choice in Lake Malawi cichlid species, but the morphological basis for this is not understood (Plenderleith et al. 2005; Takahashi and Koblmüller, 2011). Beyond this, few studies have quantitatively measured selection acting on individual phenotypes in cichlids. Disruptive selection on traits is instead typically inferred through comparisons of trait magnitude within and among diversifying lineages (Kingsolver et al. 2001).

1.4 The lateral line system

1.4.1 "Touch at a distance"

The lateral line system is an important mechanosensory network found in all fish (Figure 1.3; Figure 1.4) (Webb 2014). It provides information about habitat flow regimes (Montgomery et al. 1997; Kulpa et al. 2015), proximity to substrate, and the presence and behaviour of other organisms (Stewart et al. 2014; Butler and Maruska, 2015; Mogdans 2019). It is an ancient and unique sensory modality, distinct from (but related to) hearing and touch, and whilst ubiquitous in aquatic environments, it is unlike anything found in fully terrestrial animals (Coombs et al. 1989; Gelman et al. 2007). Because of this, it can be difficult to comprehend and has been often misunderstood. Quite aptly referred to as "touch at a distance" (Dijkgraaf and Kalmijn, 1963), it allows organisms to detect mechanical stimuli without physical contact with the source. Detection is instead achieved through sensing the movement of water around them. For example, fish use their lateral line systems to detect organisms moving in the water by sensing the water they displace. Through this they can gain information about the nature of the organism generating the movement and respond accordingly (Webb 2014; Mogdans 2019). The lateral line system is composed of a

network of functional organs called neuromasts which are located around the head, body and fins, and most strikingly along the lateral side trunk of the fish, often forming a distinct line, giving the system its name. These neuromasts serve to transduce hydrodynamic stimuli into electrical impulses, providing the organism with vital information about their immediate environment (Bleckmann and Zelick, 2009; Webb 2014).

It has been suggested that the lateral line system is a primitive vertebrate character, which evolved early in vertebrate evolutionary history. This is supported by its ubiquity across vertebrates: it is present in all fishes, including lampreys (Gelman et al. 2007), hagfishes (Braun and Northcutt, 1997), lungfishes (Northcutt 1986), coelacanths (Hensel and Balon, 2001), and all teleosts (Webb 2014) (Figure 1.3; Figure 1.4). Most amphibians in their larval forms possess a lateral line system, which some fully-aquatic amphibian adults retain (Lannoo 1988; Webb 2014) (Figure 1.3). In addition to this, studies have concluded that the system is important for fishes in numerous behaviours (Montgomery et al. 2014; Mogdans 2019), many of which are vital for survival, which implies that any ancestral aquatic vertebrate will have been under extremely strong selection pressures to develop the ability to detect flow stimuli. A close relationship of the lateral line system to other vertebrate mechanosensory systems has been proposed in several studies. Specifically, the hair cell as a signal transducer is common to the lateral line, electroreceptive, vestibular and auditory systems of vertebrates, implying a common origin (Fritzsch and Straka, 2014; England and Robert, 2022). In addition to cellular anatomy and physiology, the ontogeny of these systems and the types of stimuli involved also suggest they have shared ancestry (Webb 1989; Baker et al. 2013; Ladich and Schulz-Mirbach, 2016).

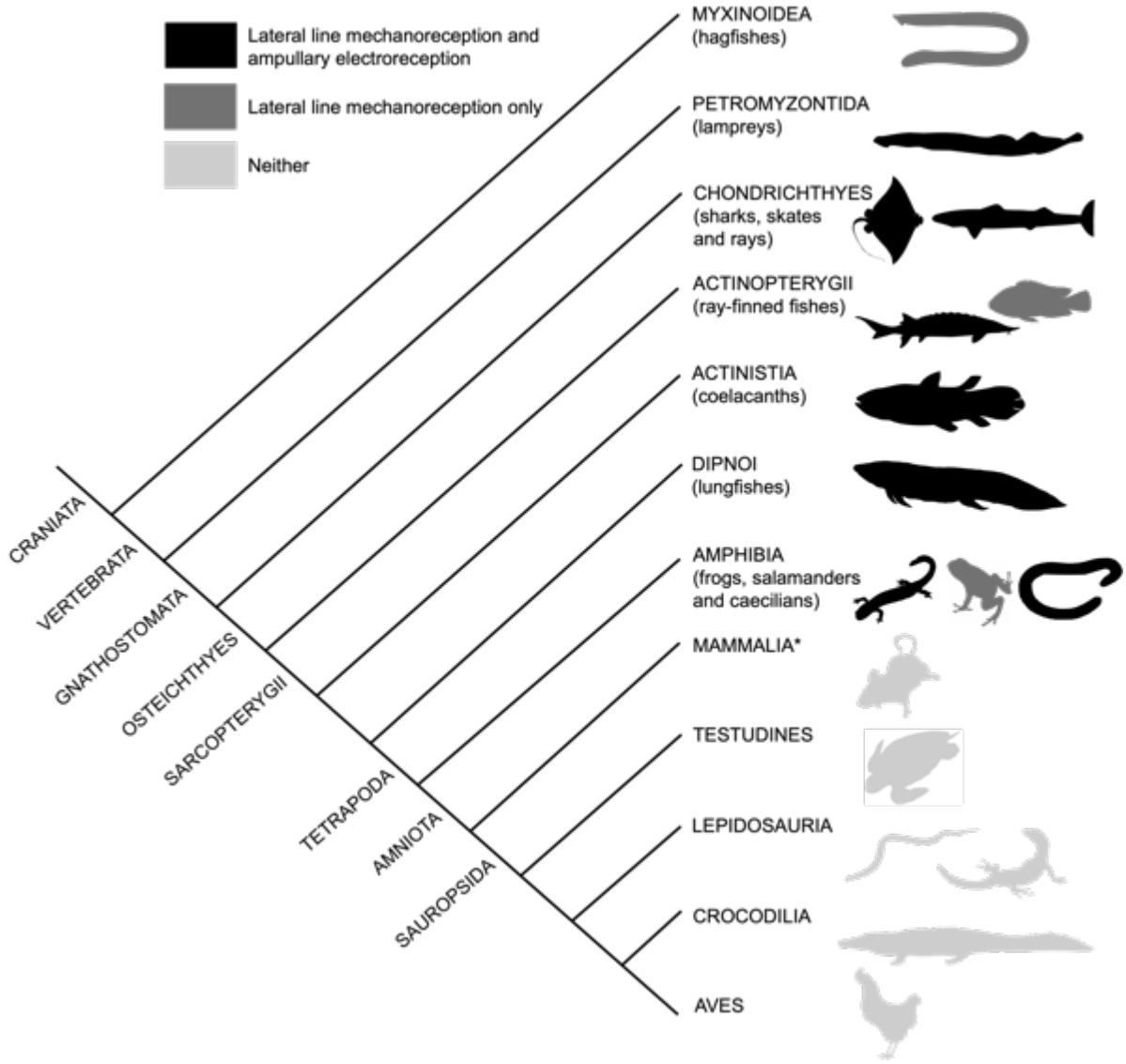


Figure 1.3 Cladogram of extant craniates. Lineages possessing lateral-line mediated mechanoreception and ampullary passive electroreceptionare highlighted in grey and black. Some clades within the Actinopterygii have evolved other forms of passive (non-ampullary) electroreception or active electroreception and these are not shown here. Lateral line system mechanoreception and ampullary electroreception are shown as present in amphibians, though in some species these are exclusively juvenile characters. Figure is adapted from Wulliman & Grothe (2014). *Electroreception present in some clades having evolved independently, and not derived from the lateral line system (within Monotremata and Cetacea, the Guiana dolphin (Czech-Damal et al. 2011), the duck-billed platypus (Gregory et al. 1987) and the Western long-beaked echidna (Gregory et al. 1989)).

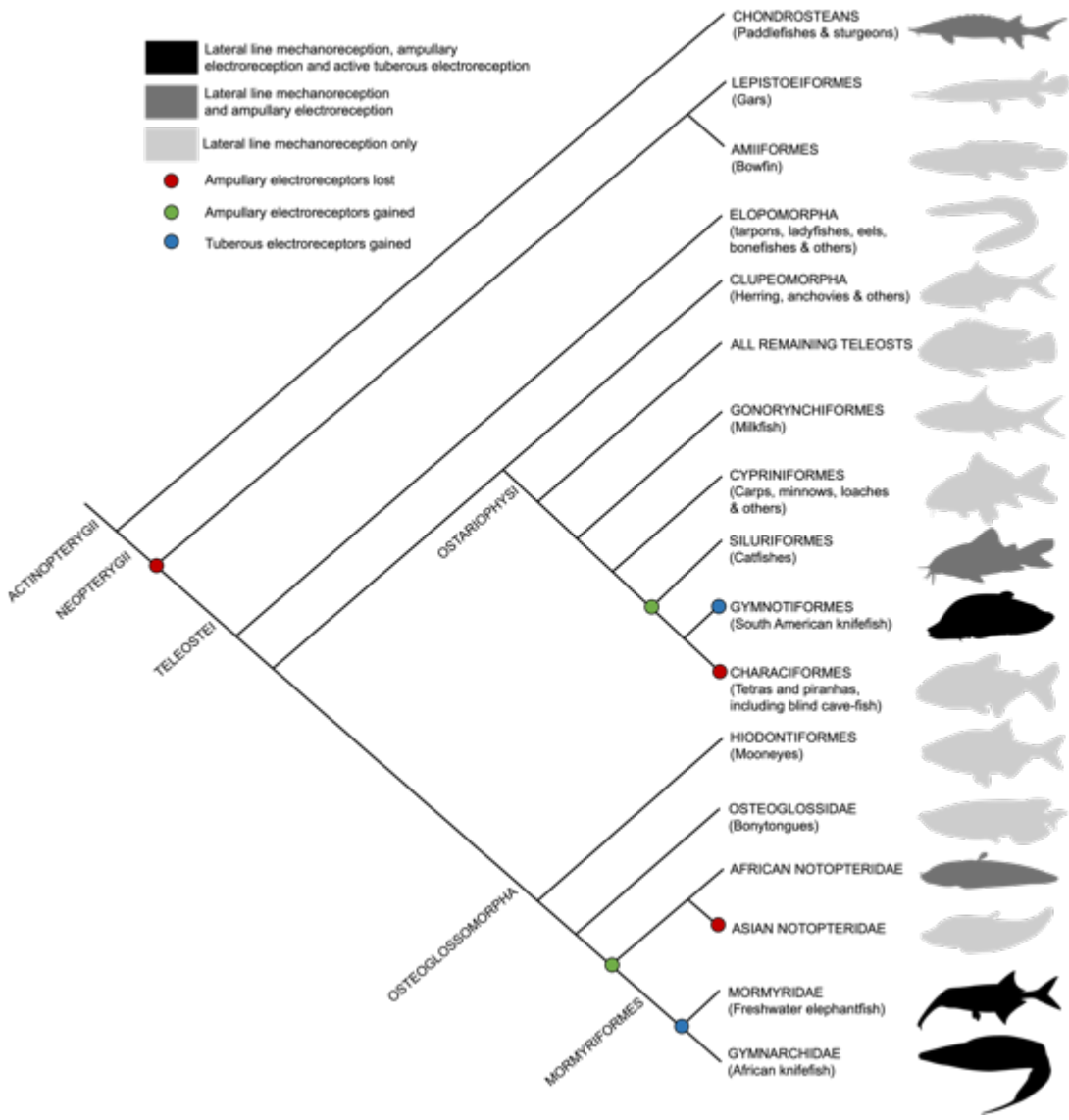


Figure 1.4 Cladogram of extant Actinopterygii showing those possessing lateral line mechanoreception, secondarily gained ampullary electroreception, and tuberous active electroreception. Adapted from Baker et al. (2013); and Wulliman & Groethe (2014). Illustrated is the hypothesised two independent instances of the evolution of ampullary electroreception in the teleosts (in the Osteoglossomorphs and the Ostariophysans, highlighted in green). In addition, active electroreception through tuberous electroreceptors is thought to have evolved independently twice, once in the old world Mormyriformes and once in the new world Gymnotiformes, highlighted in red (Near et al. 2012; Baker et al. 2013).

Study of the lateral line system has gathered interest in multiple scientific fields. It has been used as a model system for vertebrate hair cells, such as those of the auditory and vestibular systems. In medicine, for example, research into the regeneration of lamprey lateral line hair cells has been used to inform treatments for hearing loss (Jiang et al. 2014). In addition, it is of interest in drug development: neuromasts have been studied as a pre-clinical indicator of potential drug ototoxicity (Chiu et al. 2008; Buck et al. 2012). It has also been used to inspire development of biomimetic flow sensors (artificial lateral lines) which may have applications in engineering, including for autonomous underwater vehicles (Liu et al. 2016). In the biological world, it has been of interest to sensory ecologists enabling insights into the physiological and biophysical mechanisms of mechanotransduction (Montgomery et al. 1994; van Netten and McHenry, 2014), and for evolutionary/behavioural ecologists investigating its role in enabling survivorship and promoting large-scale adaptive radiation (Montgomery et al. 1994; Edgley and Genner, 2019).

1.4.2 The neuromast

The lateral line system is comprised of a network of sensory organs known as neuromasts, of which there are two types: superficial and canal neuromasts. Superficial neuromasts are located on the skin and detect movement at the organism's boundary layer (Figure 1.5A), whereas canal neuromasts are recessed within sub-dermal canals (Figure 1.5B-C) (Webb 2014; Klein and Bleckmann 2015; Becker et al. 2016). A neuromast consists of hair cells encased within a cupula (a jelly-like structure) which projects into the medium. Deflection of the cupula in response to fluid movement in turn deflects the stereovilli and kinocilia of hair cells. This acts to transduce hydrodynamic stimuli from the organism's environment into electrical impulses along afferent neurones to the central nervous system. The cupula's deflection causes changes to the membrane potential of the hair cell. When the membrane depolarises, glutamate (a neurotransmitter) is released into the synapse between the hair cell and the afferent neurones innervating it. Changes in cupula deflection cause changes in depolarisation rate, which in turn causes changes in the firing rate of action potentials along its afferent neurones (Montgomery et al. 2014; Webb 2014).

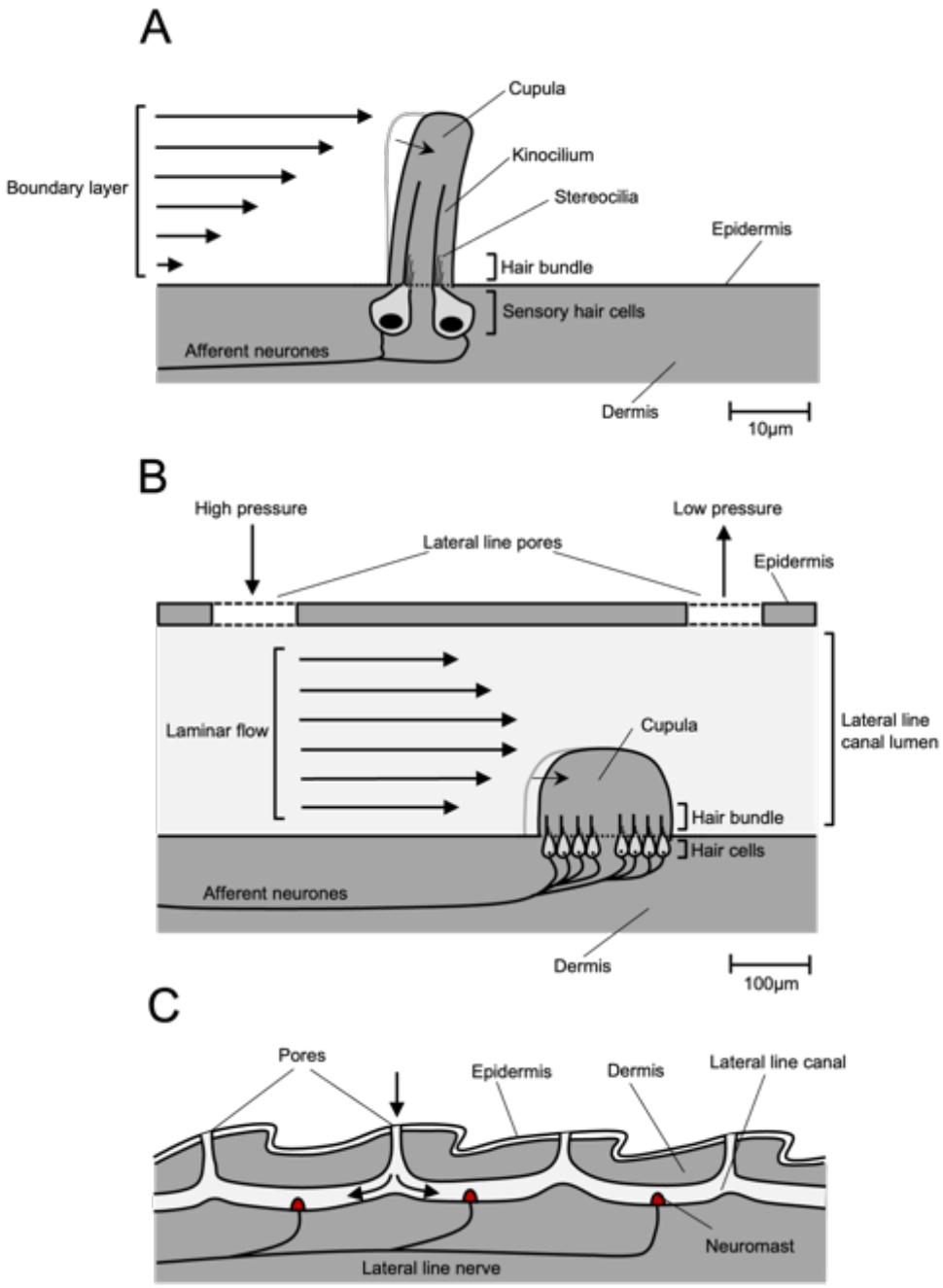


Figure 1.5 The structure of the primary sensory organ of the lateral line system: the neuromast. (A) Generalised superficial neuromast structure, located on the boundary layer of the fish. (B) Generalised structure of a canal neuromast, detecting laminar flow through sub-dermal canals. (C) The structure of a lateral line canal, showing the location of canal neuromasts between adjacent pores. Figures partially adapted from Windsor and McHenry (2009) and McHenry and van Netten (2007). Scales are approximate and highly dependent on species morphology and size.

1.4.3 Superficial and canal neuromasts

In superficial neuromasts, which are present on the epidermal layer of the organism, cupulae project into the external environment and sense movement at the epidermal layer (Figure 1.5A). In contrast, canal neuromasts project into the lumen of lateral line canals, and detect the laminar

flow of fluid through these canals, driven by differential pressure between adjacent pores (Figure 1.5B-C). The major anatomical features of both types of neuromast are the same, though the kinocilia of the hair cells in canal neuromasts do not serve a function distinct from the rest of the hair bundle, as is the case for superficial neuromasts (Laurà et al. 2018; Montgomery et al. 2014). As well as the differences between the two types of neuromast, there is also anatomical variation within the two types of neuromast. Differences in neuromast anatomy (along with other aspects of lateral line system anatomy) are thought to be related to differences in response properties, and as such the types of stimuli they can effectively detect (Webb 2014).

The different structure and arrangement of superficial and canal neuromasts is related to their different mechanoreceptive abilities, and hence functions (Montgomery et al. 2014; Webb 2014; Becker et al. 2016). The location of superficial neuromasts on the epidermis means they can detect the direction and velocity of water flowing over the surface at the boundary layer (Weeg and Bass, 2002) (Figure 1.5A). In contrast, the recession of canal neuromasts into canals allows for a more complex mechanoreception (Figure 1.5B-C). As fluid flows past canal pores, a pressure differential forms between adjacent pores forcing fluid through the canal. This results in a deflection of the cupula of the canal neuromast in the direction of the moving fluid (Peach and Rouse, 2000; Webb 2014). In this way, the organisms can compute a directional response to movement of current. In addition, canal neuromasts can detect the acceleration of water, and filter out potentially significant background hydrodynamic noise (Fritzsch and Straka, 2014). The combination of the two systems allows for a complex and adaptive mechanosensory network allowing fish to respond to a range of stimuli in differential hydrodynamic regimes (Webb 2014; Montgomery et al. 2014).

Superficial neuromasts tend to respond better to lower frequency (>30Hz) direct-current flows, while canal neuromasts respond better to high-frequency alternating-current flows (~30 to 150Hz), thus acting as low-pass and high-pass filters respectively (Coombs et al. 2001; Klein and Bleckmann, 2015). In part due to these differential response properties, superficial and canal neuromasts differ in how efficiently they respond to stimuli. Superficial neuromasts tend to respond most effectively to abiotic stimuli (i.e. contributing to the rheotactic response to the bulk movement of water) (Weeg and Bass, 2002), whereas canal neuromasts respond most effectively to stimuli that are punctuated and biotic in nature (i.e. responding to the movement of another organism) (Coombs et al. 2001; Klein and Bleckmann, 2015). Considering these detection capabilities, along with their general anatomical placement, superficial neuromasts can be considered to primarily detect uniform flow velocity, whether laminar or turbulent (Montgomery

et al. 1997). Superficial neuromasts tend to be most numerous in species living in habitats with slower-moving water (Wark and Peichel, 2010). In contrast, canal neuromasts are capable of responding to pressure differentials and fluid acceleration, and are more likely to be dominant in less uniform flow, for example those generated by a swimming prey item (Coombs et al. 2001; Engelmann et al. 2002; van Netten 2006; McHenry and Liao 2014; Mogdans 2019). It is thought that canal neuromasts are particularly important for those fish in environments with significant background hydrodynamic noise, such as high-flow riverine systems (Wark and Peichel, 2010), where high-pass filters will be most valuable (Coombs et al. 2001). Even though these trends can be found among many teleost groups, the complexity of the system's anatomy and the large range of behaviours it is used for means that it often manifests in unpredictable ways, not necessarily adhering to the general patterns described here.

1.4.4 Morphological diversity of the lateral line system

The layout of lateral line canals and the number and distribution of both types of neuromast is highly variable, even among teleosts. It tends to comprise a network of canals containing canal neuromasts, and superficial neuromasts on the epidermis, on scales, or in pits or grooves (Webb 2014). The term "lateral line" refers to the trunk canal, which in most cases runs straight along the organism's flank from near the posterior limit of the operculum to near the caudal peduncle. Despite this name, the canals of the lateral line system can also be found in the head: around the eyes, and under the jaw through the preopercular bone and angulo-articular bones, and into the dentary bone (Figure 1.6).

Number and patterning of superficial neuromasts varies significantly across teleost species. In some species they are found in pits or grooves in the fish epidermis, and in others, particularly deep-marine species, they can be situated on projections or processes holding them above the epidermis surface (stalks or filaments) (Webb et al. 2014; Marranzino and Webb, 2018). Some species have evolved various accessory structures around their superficial neuromasts (Webb 2014). On the level of the individual neuromast, there is morphological diversity beyond the differences between superficial and canal neuromasts, in size and shape. Much of this morphological variation is thought to be adaptive, and in most cases is likely to have developed in response to hydrodynamic selection pressures both biotic and abiotic in nature (Wark and Peichel, 2010).

The cephalic canals

In general, there are up to seven lateral line canals in the fish head. These are the supraorbital canal (within the frontal and nasal bones), the infraorbital canal (within the lacrimal bone and other associated structures), the preopercular canal (within the preopercular bone), the mandibular canal (within the angulo-articular and dentary bones), the otic canal (within the pterotic bone), the post-otic canal (within the posttemporal bone) and the supratemporal canal (within the medial and lateral extrascapular bones) (Tarby and Webb, 2003; Webb 2014) (Figure 1.6). The seven cephalic canals are not all present in every species; canals may be enlarged, reduced or lost altogether, in ways typically related to ecological or behavioural selection pressures (Wark and Peichel, 2010; Schwalbe et al. 2012; Webb 2014; Klein and Bleckmann, 2015).

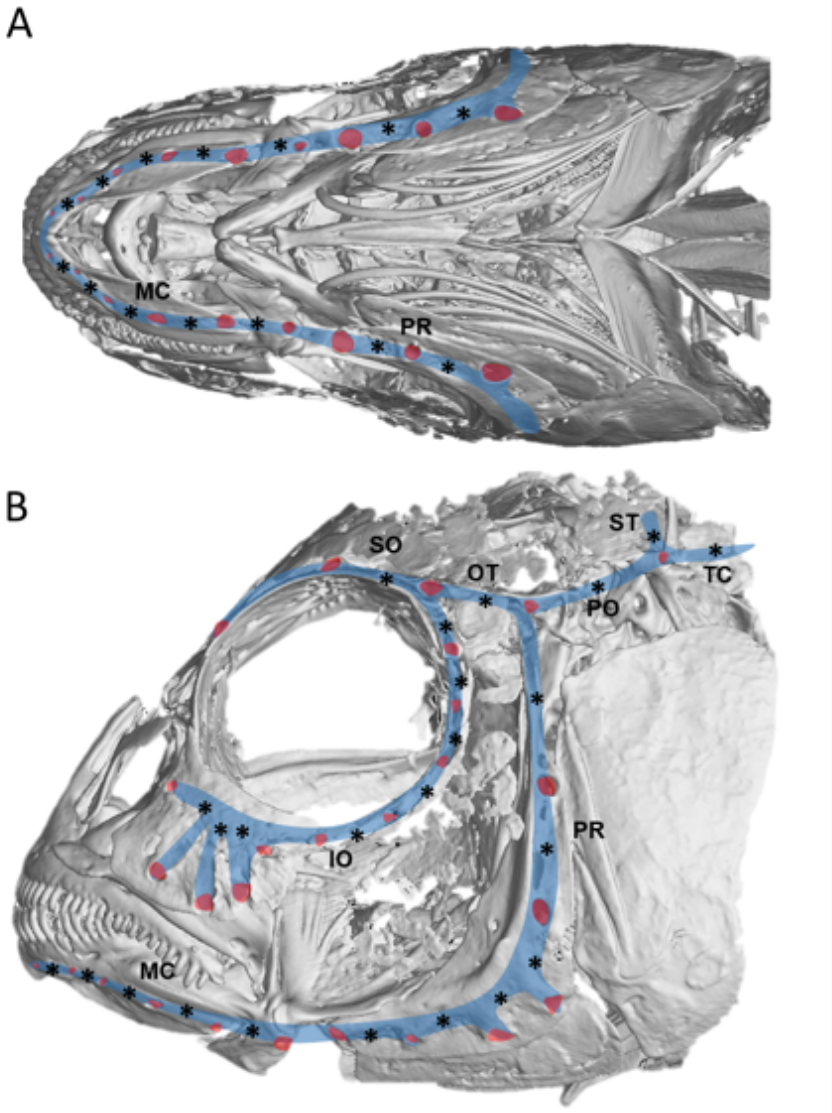


Figure 1.6 Overview of the cephalic canal lateral line system of the Lake Malawi cichlid *Astatotilapia calliptera* (A) Ventral view, and (B) lateral view. The positions of canals (blue) and pores (red) are shown. Approximate position of canal neuromasts are shown as asterisks, as determined by DASPEI staining. MC = mandibular canal; PR = preopercular canal; SO = supraorbital canal; IO = infraorbital canal; OT = otic canal; PO = post-otic canal; ST = supratemporal canal; TC = trunk canal.

Positioning of canal neuromasts is ascertained from staining of multiple *Astatotilapia calliptera* individuals with DASPEI. Figure is adapted from Edgley & Genner (2019).

It has been argued that across teleost groups the cephalic canal lateral line system diversity can be categorised into five broad groups: simple narrow canals, branched narrow canals, narrow canals with widened tubules, reduced canals, and widened canals (Webb 2014) (Figure 1.7). Narrow cephalic lateral line canals are extremely common across the teleosts (Figure 1.7A), whereas a widened cephalic canal phenotype has been thought to be relatively uncommon (Coombs et al. 1989). It is found in deep-water marine species but seems to have also evolved convergently in several freshwater taxa (Coombs et al. 1989; Webb 2014). These canals are thought to be a sensory adaptation for detection of prey (Schwalbe et al. 2012), and it has been found that these large pores and wide canals increase sensitivity to punctuated stimuli such as those made by potential prey items (Schwalbe and Webb, 2014). These enlarged pore phenotypes are particularly pronounced on the anterior and ventral canals and pores of the head (Edgley and Genner, 2019). This is supported by the fact that the trunk canal tends not to exhibit a widened phenotype in any species (Webb 2014), and behavioural trials showing increased prey-detection capacity in species with widened canals in dark conditions (Schwalbe et al. 2012) (Figure 1.7B).

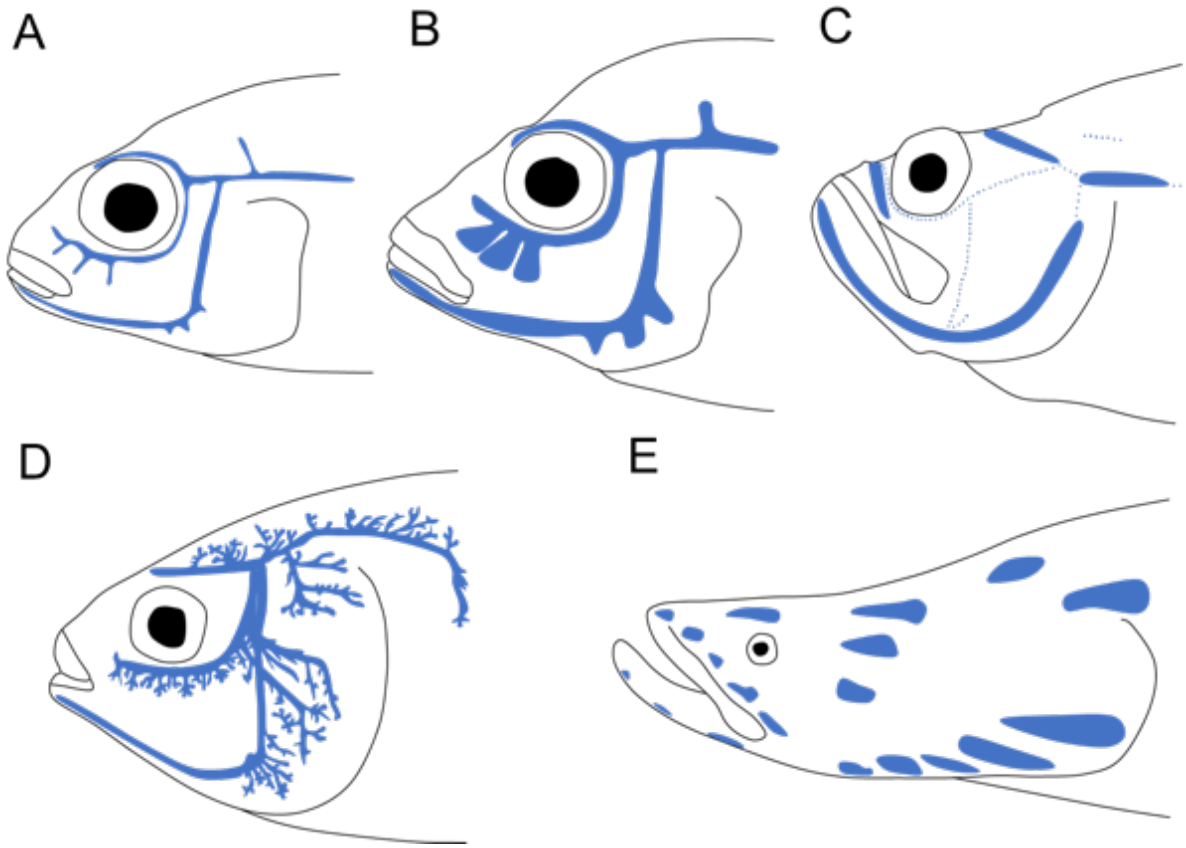


Figure 1.7 The main categories of cephalic canal systems in teleosts. Adapted from Webb (2014) (A) Simple narrow canal system of the Eastern happy (*Astatotilapia calliptera*); (B) Simple widened canal system of Grant's peacock cichlid (*Aulonocara stuartgranti*); (C) Reduced canal system accompanied with lines of superficial neuromasts, on the plainfin midshipman (*Porichthys*

notatus); (D) Narrow canal system with branched canals of the Atlantic menhaden (*Brevoortia tyrannus*). Each bony pore opens into multiple highly branched tubules leading to many terminal pores; (E) Narrow canals with widened tubules illustrated on *Arapaima gigas*. Narrow canals lead to 'pouch-like' epithelial tubules.

Some species have lost their cephalic lateral line canals altogether: some populations of three-spined sticklebacks completely lack canal neuromasts (Wark and Peichel, 2010). Several reasons have been suggested for this. It may due to paedomorphosis, an adaptation for nocturnal hunting, or as an adaptation for slow-moving water. All Tetraodontiformes (which includes the boxfishes and pufferfishes) have lost their lateral line canals completely, perhaps related to the unique and radical departures from usual fish anatomy present within the order (Nakae and Sasaki, 2010; Li et al. 2018). The "widened tubules" phenotype of *Arapaima* and others is also relatively uncommon, as it appears to be used for a unique purpose (Figure 1.7E). In *Arapaima gigas*, a whitish fluid is secreted from the integument of the cavities into the surrounding water. This has been suggested to be related to the immune response, or perhaps parental care given that the frequency of release increases post-fertilisation (Torati et al. 2017).

The trunk canal

There is significant diversity in trunk canal morphology among teleosts. These can be broadly categorised into eight morphological groups (Figure 1.8), although most teleosts have a single, relatively straight and unbroken trunk canal (Figure 1.8A). A few species with particularly distinctive habits have evolved trunk canals related to these functions. For example, some benthic species have developed a single, dorsally located trunk line (Figure 1.8C) (Montgomery and Coombs, 1998), whereas other surface-associated species (such as the flying fishes of the family Exocoetidae) have evolved a single ventrally-located trunk canal (Figure 1.8D) (Parin and Astakhov, 1982). It is thought that this is due to the importance of detecting hydrodynamic stimuli from above and below respectively (Webb 2014). In some species the trunk canal has been lost altogether, commonly in schooling and planktivorous fishes (Webb 1989) (Figure 1.8H). Cichlids, such as the haplochromines of the Lake Malawi radiation, mostly have a disjunct trunk canal (Figure 1.8F), though there is some variation. The reason for the disjunct canal phenotype is not entirely known (Webb 1990), though it may be related to minimising the effect of self-generated flow from the pectoral fins. Functionally, the reason for variation in trunk canal patterning within teleosts is not well understood.

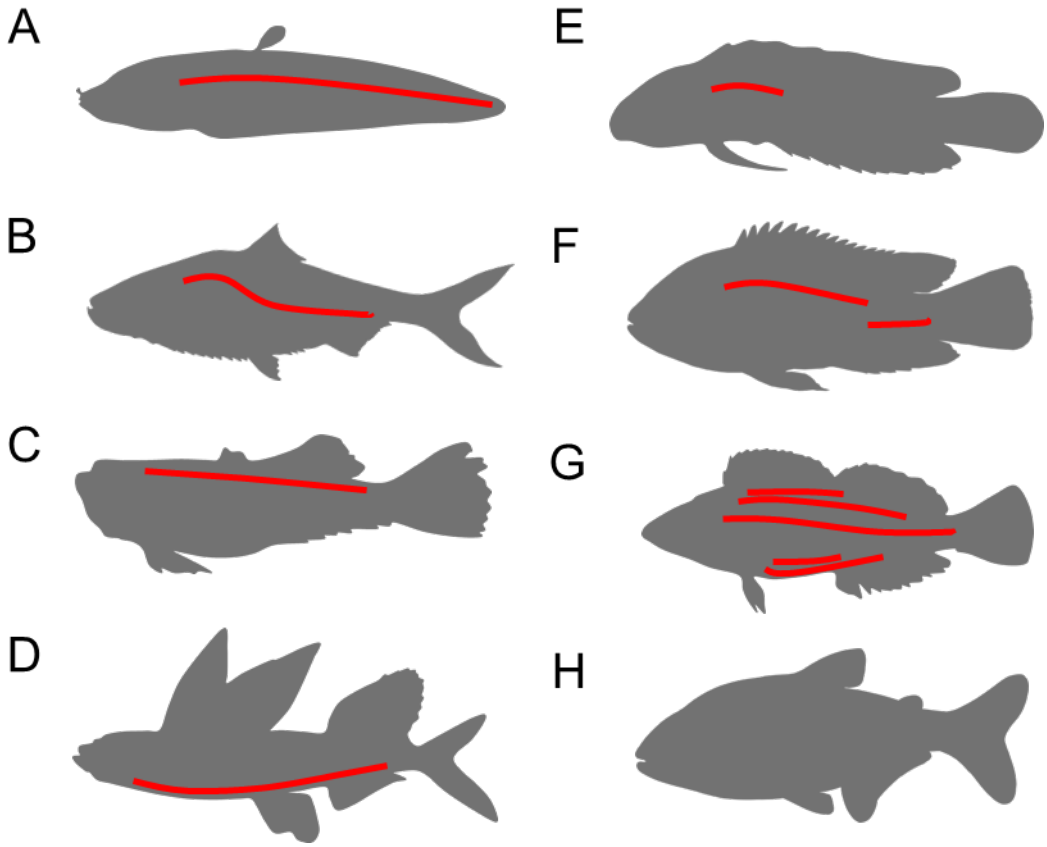


Figure 1.8 The diversity in trunk canal morphology in teleosts. Adapted from Webb, 2014. (A) Complete straight trunk canal; (B) Complete arched trunk canal; (C) Dorsally oriented trunk canal; (D) Ventrally oriented trunk canal; (E) Incomplete trunk canal; (F) Disjunct trunk canal; (G) Multiple trunk canals; (H) Absent trunk canal.

Superficial neuromast patterning

The distribution and total number of superficial neuromasts across the body varies both within and between species. The location of neuromasts can be determined by staining with DASPEI ((2-(4-(Dimethylamino)styryl)-N-ethylpyridinium iodide), a vital mitochondrial dye, followed by imaging under a fluorescent lamp and filter (Buck et al. 2012). Superficial neuromasts tend to be grouped in clusters or arranged in lines, often associated with lateral line canals and canal neuromasts. In cichlids, they are generally arranged around canals and canal pores in clusters, and along the caudal fin in rows. Quite often when canals are absent, there is a line of superficial neuromasts in their place (Webb 2014). In cichlids, which possess disjunct canals, the anterior trunk canal segment seems to be completed by a line of superficial neuromasts which stop at the caudal peduncle. This is most likely related to the development of the disjunct canal in cichlids and the role of rate heterochrony in the process (Webb 1990; Bird and Webb, 2014). Patterning of superficial neuromasts has been found to vary significantly within species as well. For example, absolute number of superficial neuromasts was found to vary between habitats with different flow

regimes in three-spined sticklebacks. This suggests that different selection pressures in different habitats are driving variation in superficial neuromasts within species (Wark and Peichel 2010).

1.4.5 Relating morphological diversity to ecology

Interspecific and intraspecific variation in lateral line system structure has been quantified in cichlids (Webb 2014), sticklebacks (Wark and Peichel, 2010; Wark et al. 2012), notothenioid icefishes (Montgomery et al. 1994), Trinidadian guppies (Fischer et al. 2013), and other fish lineages (e.g. Wellenreuther et al. 2010). This variation includes neuromast arrangement (Trokovic et al. 2011; Schwalbe and Webb, 2014), number (Wark and Peichel, 2010), size (Coombs and Montgomery 1994), and type (Coombs et al. 2001; Schmitz et al. 2014), as well as aspects of canal architecture, like pore size, canal width and degree of bone ossification (Montgomery et al. 1994; Wellenreuther et al. 2010; Klein et al. 2013). Patterns of diversity in morphology are often adaptive and related to specific function, in particular species reliance on flow stimuli. However, due to the complexity of the system, it is not known fully how these patterns manifest in different environments. Despite this, some general trends in morphology and ecology can be determined.

In general, an increased number of both types of neuromast is related to a higher importance of flow stimuli in the environment (Bleckmann and Zelick, 2009; Kulpa et al. 2015). This has also been shown to be the case for both a widened lateral line canal phenotype (Schwalbe et al. 2012) and increased neuromast size (Webb 2014). As expected, in the environments in which flow stimuli tend to be more complex, and in which their detection is more imperative for survival, we tend to find organisms with more neuromasts. These include benthic habitats where light levels are low and/or turbidity is high (Trokovic et al. 2011; Schwalbe and Webb, 2015), and riverine habitats, in which the flow rate is generally higher and the flow more turbulent than lacustrine or marine environments (Wark and Peichel, 2010). It is commonly found that lateral line system mechanoreception is or particular importance when light levels are reduced or absent, such as in cave-dwellers (Abdel-Latif et al. 1990; Kulpa et al. 2015) or deep-water species (Schwalbe and Webb, 2015; Marranzino and Webb, 2018).

The proportion and relative importance of the canal and superficial neuromasts can vary among species. A higher number of superficial neuromasts increases sensitivity to water flow, but evolving an increased number of superficial neuromasts may not be an effective way of increasing flow sensitivity above a certain threshold in environments with significant hydrodynamic noise, such as in a complex and fast flowing stream. Due to this, organisms in still or slower flowing

environments tend to exhibit proliferated superficial neuromasts and lose canal neuromasts (Engelmann et al. 2002; Wark and Peichel, 2010). Some three-spined stickleback populations, for example, have lost their canal neuromasts entirely (Wark and Peichel, 2010), and in other species with similar habits, they are highly reduced (Trokovic et al. 2011). Canal neuromasts seem to be more important in species where the water flow is more turbulent, and the ability to filter out important stimuli from background noise is of greater importance (Bassett et al. 2006). This tends to manifest as a greater number of canal neuromasts, and wider canals with greater ossification (Bassett et al. 2006; Schwalbe and Webb, 2014). This is not a hard and fast rule, due to the lateral line system's importance in a multitude of behaviours, and unpredictable habitat hydrodynamics, but in general, many slower swimmers in still waters have predominantly superficial neuromasts, and active, fast swimmers in turbulent waters have a greater focus on canal neuromasts, as well as more neuromasts overall (Montgomery et al. 1995). In general, species endemic to marine habitats tend to show a lower number of both canal and superficial neuromasts, perhaps due to lower variation in flow regimes. As well as this, marine species tend to have lower lateral line system structure heterogeneity compared to their lacustrine or riverine counterparts (Trokovic et al. 2011), perhaps due to a relative higher predation risk driving stabilising selection.

Notably, in many species neuromast phenotype is better predicted by behavioural factors than abiotic habitat hydrodynamics. In two species of the Pinguipedidae (sandperches), diet seems to be a strong selection pressure, and it seems that feeding habits (or hydrodynamic stimuli involved in feeding) influence neuromast arrangement and canal architecture more than the abiotic hydrodynamics of species environment (Carton and Montgomery, 2004). Jaw shape is highly related to trophic ecology of fishes (Albertson et al. 2003; Hulsey and García De León, 2005). Many algivorous fish species for example tend to have a flattened jaw for scraping rocks, while pelagic predatory fish tend to have laterally compressed and anteroposteriorly elongated jaws. The morphology of the cranial lateral line canals will therefore be constrained by the size and shape of some components of the oral jaw, such as the dentary and preopercular bones (which contain the mandibular and preopercular canals respectively) (Edgley and Genner, 2019). Therefore, adaptation to develop a more sensitive lateral line system may be constrained by comparatively stronger selection pressures for certain jaw morphologies, causing different lateral line canal phenotypes to those expected by specific species ecology. In addition, changes in the lateral line system may not be adaptive. Some species of Antarctic icefishes (Notothenioidei) have lost their lateral line canal system entirely, but this does not appear to be primarily due to selection pressures on the lateral line system. It has been found that paedomorphism, or the retention of some juvenile

characteristics into adulthood, has caused the reduced (or lost) canal phenotype (Montgomery and Clements, 2000).

As well as abiotic components of habitat such as turbidity, light level and water turbulence, biotic aspects of species ecology can influence structure and diversity of lateral line system structure. In the Trinidadian guppy it has been demonstrated that there is significant intraspecific variation in neuromast arrangement and number, which is dependent on predator density and hence predation risk. This not only suggests a role for the lateral line system in predator avoidance, but also suggests that there is plasticity in the development of the lateral line system, and/or that the system may be subject to rapid microevolutionary changes (Fischer et al. 2013). Additionally, diet has been linked to both neuromast (Schwalbe and Webb, 2014) and cranial canal (Edgley and Genner, 2019) morphology in cichlids, implying that trophic ecology confers selection pressures on the lateral line system. Despite this, an association between habitat or ecology and neuromast structure/distribution is not always observed, and in some taxa, there seems to be no link (Beckmann et al. 2010), perhaps suggesting that the lateral line system is less imperative for the survival of some species (Bleckmann and Zelick, 2009), or that other stronger selection pressures can dominate under certain evolutionary scenarios.

In some cases, there is a relationship between lateral line system structure and body size, head size and/or brain morphology (Montgomery et al. 1994; Bird and Webb, 2014; Webb 2014) both among and within species. For example, in notothenioid icefishes, the comparatively large *Dissostichus mawsoni* has relatively narrow lateral line canals when compared to several other smaller sympatric species. This has been suggested to be an adaptation to compensate for its larger size, with the narrowing of the canal providing the resistance needed for canal neuromasts to detect the full range of frequencies, at the cost of some sensitivity. As such, the response properties of the system are the same as in the smaller species despite the body size and canal morphology differences (Montgomery et al. 1994; Engelmann et al. 2002). It seems that this narrowing of the lateral line canals is an adaptation to improve the signal-to-noise ratio in more turbid, turbulent or unpredictable environments. Conversely, in other species, it has been suggested that canal widening, wider spacing of pores leading to the lateral line canals, and higher levels of branching in these canals are also adaptations to increase the sensitivity of the system (Klein et al. 2013).

Studies on the response properties of different lateral line system phenotypes, and their significance for the animal's ecology and behaviour, often require simplification to model the flow

around the fish or inside the lateral line canals (Klein et al. 2013), although more recently protocols have attempted to remedy this (Kerr et al. 2016). Often experiments use artificial neuromasts – biomimetic flow sensors – to model the structures of the system (Klein et al. 2013). From this work, certain general patterns have become more firmly established, including that the presence of canals can increase sensitivity of the system, and narrowing of the canals can filter out noise from signal (Windsor and McHenry, 2009).

1.4.6 The lateral line system's role in mediating behavioural responses

The lateral line system is an extremely important sensory modality in a multitude of key fish behaviours including rheotaxis (Montgomery et al. 1997), prey detection (Montgomery and MacDonald, 1987; Janssen 1996; Pohlmann et al. 2004; Schwalbe et al. 2012), predator avoidance (Stewart et al. 2014), shoaling behaviour (Faucher et al. 2010) and male-male competition (Butler and Maruska, 2016).

Rheotaxis

Rheotaxis is the alignment of a fish in response to the flow of its surroundings. The term is often used to describe alignment to flow, in which a fish adjusts its position to face the current to reduce hydrodynamic drag (Coombs et al. 2020). In doing this, the fish can maintain its position or swim upstream, whilst avoiding being swept downstream. This behaviour is found in most aquatic vertebrates (Coombs et al. 2020). Despite this, rheotaxis can be used to describe alignment to any kind of flow stimulus, abiotic or biotic in origin. Alignment towards flow is known as positive rheotaxis, and alignment away from flow is known as negative rheotaxis. The lateral line system is known to be of paramount importance for mediation of rheotactic behaviours (Montgomery et al. 1997; Kulpa et al. 2015; Coombs et al. 2020). Although vision and olfaction can also influence orientation in the water (Guthrie 1986; Kasumyan and Marusov, 2007), it is evident that the primary stimuli are hydrodynamic, as fish with their lateral line pharmacologically ablated exhibit significantly impaired rheotactic responses, particularly in the absence of visual cues (Montgomery et al. 1997; Suli et al. 2012; Kulpa et al. 2015). Superficial neuromasts have been found to be involved in the mediation of this response in fishes (Montgomery et al. 1997). Variation in lateral line morphology has been associated with differential rheotaxis performance in tadpoles: a species with a greater number of larger superficial neuromasts (*Xenopus laevis*) had a stronger rheotactic response than a species with fewer and smaller neuromasts (*Rana catesbeiana*) (Brown and Simmons, 2016). It has also been found that in flow regimes that are more complex, lateral line ablation severely impairs the rheotactic response in blind cavefish *Astyanax mexicanus* (Kulpa et al. 2015),

although the lateral line system is not necessary for rheotaxis in stiller waters (van Trump and McHenry, 2013).

Prey detection

Positive rheotaxis mediated by the lateral line system is involved in prey detection, not only in the water column (Coombs et al. 2001) but at the surface (Janssen 1996) and beneath the substrate (Schwalbe et al. 2012). Although vision (Guthrie 1986), olfaction (Hara 1986) and electroreception (Kalmijn 1966) are all utilised in finding and moving towards prey, the lateral line system has been found to also mediate both the initial detection of (Montgomery and MacDonald, 1987) and the alignment towards (Coombs et al. 2001; Pohlmann et al. 2004) prey items in several species. It has been found to be particularly important for this response in dark conditions (Pohlmann et al. 2004; Schwalbe et al. 2012). Populations of blind cavefish are thought to have evolved sensitive lateral line structures related to prey detection and capture. Being unable to use vision like their non-cave-dwelling relatives, they have developed a lateral-line mediated prey capture strategy (Yoshizawa et al. 2010; Yoshizawa et al. 2014; Lloyd et al. 2018). As well as mediating the orienting response towards prey, the lateral line system is known to be used while actively searching for prey items. Some cichlid species often referred to as 'sonar feeders' probe the substrate with the widened canals of their head in search of the movement of buried prey (Turner, 1996). In the inverse direction, the stargazer *Leptoscopus macropygus* has been observed sitting stationary on the substrate, waiting to detect the flows of a passing prey item with its unique dorsal-facing lateral line canals (Montgomery and Coombs, 1998).

The role of the lateral line system in detection of prey is highly dependent on the nature of the prey items involved, as this affects the hydrodynamic stimuli they generate (Montgomery et al. 1994). For example, surface feeding fishes preying on insects respond to relatively conserved and hence predictable hydrodynamic stimuli (Müller and Schwartz, 1982), whereas larger organisms in the water column, such as a swimming fish, create flow patterns in their wake that are more complex and harder to predict (Bleckmann et al. 1991; Pohlmann et al. 2004), shedding vortices and creating turbulence (McHenry and Liao 2014; Montgomery et al. 2014; Kulpa et al. 2015).

Predator avoidance

The avoidance of predators will often involve a negative rheotaxis. This may manifest in moving rapidly away from an approaching individual perceived as being a potential threat, or responding to the strike of a predator (Montgomery et al. 1995; Higham et al. 2015). In the case of the former,

choosing whether to align oneself towards or away from a flow stimulus is a critical choice for fishes, in which both detecting smaller prey items and avoiding larger predators are imperative for survival. There is a trade-off in responding to flow stimuli, as the decision must be taken extremely rapidly, as the cost of making the incorrect choice may be lethal. However, these decisions are not being made based on flow stimuli alone, as the dynamics of the rheotactic response depends on the availability of other forms of information (e.g. visual), the speed and form of the flow stimulus, and the background hydrodynamic noise (Bak-Coleman et al. 2013).

Flow detection is known to be important for the escape response of fish. The ability of zebrafish larvae (*Danio rerio*) larvae to escape from an attacking conspecific adult is almost entirely removed by pharmacological lateral line ablation (Stewart et al. 2014). Specifically, in zebrafish the lateral line system enables predator detection prior to initiation of an attacking strike, and there is a three times higher chance of evading the attack of a suction-feeding predator when a negative rheotactic response was initiated before the predator's mouth opened. Detecting and responding to potential threats before an attack was found to be primarily informed by lateral line system mechanoreception rather than visual cues, by detecting the bow wave of an approaching predator (Stewart et al. 2014).

Schooling and shoaling

Shoaling (aggregations of individuals) and schooling (coordinated and synchronised movement of individuals together) are both important behaviours for fish survival (Miller and Gerlai 2012; Herbert-Read et al. 2017). In some cases fish have been described as acting as a decentralised 'superorganism', with each individual moving in synchrony with the others by responding to cues of their nearest neighbours (Huth and Wissel, 1994; Miller and Gerlai 2012). Both vision and lateral line system mechanoreception are thought to play a role in the initiation and coordination of fish collective motion (Partridge and Pitcher, 1980; McKee et al. 2020). Some fish species can school or shoal at night, and in deep or turbid environments where vision is not reliable, implying this is somewhat mediated by their lateral line systems (Gregson and de Perera, 2007; Kowalko et al. 2013). Chemical ablation of lateral line neuromasts has been found to impair ability to form both schools (Mekdara et al. 2018) and shoals (Pitcher 1979; Partridge and Pitcher, 1980). Combined, these studies show that lateral line system mechanoreception does inform collective movement of fishes, although the detailed relationship with vision is not well understood (Partridge and Pitcher 1980).

Other social behaviours

The lateral line system is thought to be used to inform many social interactions in fishes, including those involved with spawning, mating, defending territories, and communicating with conspecifics (Montgomery et al. 2014). One such example of a lateral line-mediated social interaction has been found in cichlids, which were found to use their lateral line systems to assess opponents during aggressive interactions between males (Butler and Maruska 2015). In addition to this, a role for the lateral line system in spawning and mating behaviours have been reported, where lateral-line ablated fish have been found to have lower reproductive success in several species (Satou et al. 1994; Mirjany et al. 2011; Medina et al. 2013). As well as detection, hydrodynamic cues are thought to be produced by some fish as a means of communicating via their lateral line system (Satou et al. 1994; Medina et al. 2013), and this may be more widespread than is commonly assumed (Butler and Maruska, 2016). Despite this, there is little experimental or field-based evidence for the function of the lateral line system in mediating social interactions. This is mainly due to problems decoupling the input of other sensory systems, such as vision and olfaction, and limitations of the chemical ablation methods commonly used to infer lateral line system function (Butler et al. 2016; Mekdara et al. 2022).

Experimental Ablation of the Lateral Line System

The oxotoxin gentamicin is commonly used to ablate the lateral line system in experimental trials. It is known to cause damage to neuromasts (Montgomery et al. 2014), affecting function in both amphibians (Brown and Simmons 2016) and fishes (Buck et al. 2012). However, there is conflict in the literature over whether it affects both canal and superficial neuromasts, and whether any function is retained (Brown et al. 2011; Van Trump and McHenry, 2013; Brown and Simmons, 2016; Mekdara et al. 2022). Given this uncertainty, the use of gentamicin or other ototoxic antibiotics specifically for selective ablation of either superficial or canal neuromasts should be avoided. In addition to antibiotic ablation, several behavioural studies have relied on cobalt chloride ($COCl_2$) treatment to determine the effect of a disabled lateral line on fish behaviour (e.g. Coombs et al. 2001; Schwalbe et al. 2012; Butler and Maruska, 2015). This treatment has been shown to also impair all superficially located sensory receptors, including olfactory receptors, in a range of freshwater fish, including the African cichlid *Astatotilapia burtoni* (Butler et al. 2016). Olfaction is known in fish to be vital for both the detection of prey (Chivers and Smith, 1998) and assessing predation risk and facilitating predator avoidance (Wisenden 2000; Fischer et al. 2017). Since there is evidence that both olfaction and mechanoreception via the lateral line serve these

important functions (Schwalbe et al. 2012), when assessing the relative importance of the lateral line to behaviour, the impacts of $CoCl_2$ treatment on the olfactory system must be accounted for.

1.4.7 Limitations of lateral line system research

The adaptive function of this considerable diversity in lateral line systems, on the level of the neuromast and in terms of larger arrangement of the system, is poorly understood (Webb 2014). The structures of the lateral line system can vary in numerous ways, and so it is difficult to efficiently categorise species. The use of metrics like absolute neuromast numbers, or canal to superficial neuromast ratios, is flawed, due to differences in neuromast patterning, size and morphology also affecting individual flow-detection ability (Webb 2014). Furthermore, morphological differences in lateral lines do not often necessarily result in a differential physiological response, or act in the way predicted (Coombs and Montgomery, 1992). Due to the occurrence of substantial interspecific (Wark and Peichel 2010; Wellenreuther et al. 2010) and intraspecific morphological lateral line variation (Fischer et al. 2013), care must be taken when extrapolating limited data sets and assuming the phenotype of any species or population based on habitat or otherwise (Beckmann et al. 2010; Schwalbe and Webb, 2015). Despite this, we may expect changes to the lateral line system through adaptation to manifest in predictable ways through convergent evolution, perhaps in part due to the importance of phenomena like rate heterochrony in some aspects of phenotypic diversity, which may be common, due to the way it can facilitate fast adaptation (Bird and Webb, 2014).

1.4.8 The lateral line system and evolutionary biology – a missed opportunity

Despite the wide-ranging functions of the lateral line system in fishes, and its ubiquity, there has been relatively little research on the role of the lateral line system in speciation and adaptive evolution. For example, in radiating lineages such as cichlids, current research has focused not on assessing the diversity present, but morphological comparisons between species pairs (Becker et al. 2016). Determining the role of the lateral line system in adaptive evolution will require identifying the genetic basis of lateral line system diversity, in which several studies have gained key preliminary insights (Wark and Peichel, 2010; Wark et al. 2012; Becker et al. 2016). Research on this to date has focused only on the superficial neuromasts of the trunk, identifying candidate genes associated with superficial neuromast morphology, such as the *Eda* locus (Wark et al. 2012; Mills et al. 2014; Greenwood et al. 2016). *Eda* has been found to be involved in mediating both neuromast number and bony plate morphology in sticklebacks (Archambeault et al. 2020), and schooling behaviour (Greenwood et al. 2016). However, finding the genetic basis for the

phenotypic diversity across the varied morphology of the lateral line system will prove to be challenging, due to epistatic and pleiotropic effects of any genes which are related to such diverse behaviours and morphologies (Albertson et al. 2003; Wark et al. 2012; Mills et al. 2014; Greenwood et al. 2016; Archambeault et al. 2020).

In a recent study, the variation in cranial lateral line morphology within a fish adaptive radiation was described for the first time, focusing on 52 species of haplochromine cichlids from the Lake Malawi cichlid radiation (Edgley and Genner, 2019) (Figure 1.9). While accounting for the constraints of jaw shape, it was found that cranial lateral line system morphology was associated with several aspects of morphology, most notably the diet of each species. These findings comprised the first evidence of widespread morphological diversification in this sensory system in the adaptive radiation process, indicating that the system is under strong disruptive selection during diversification. This suggests that the system has a key role in species adaptation to new ecological challenges, particularly trophic niche (Edgley and Genner, 2019).

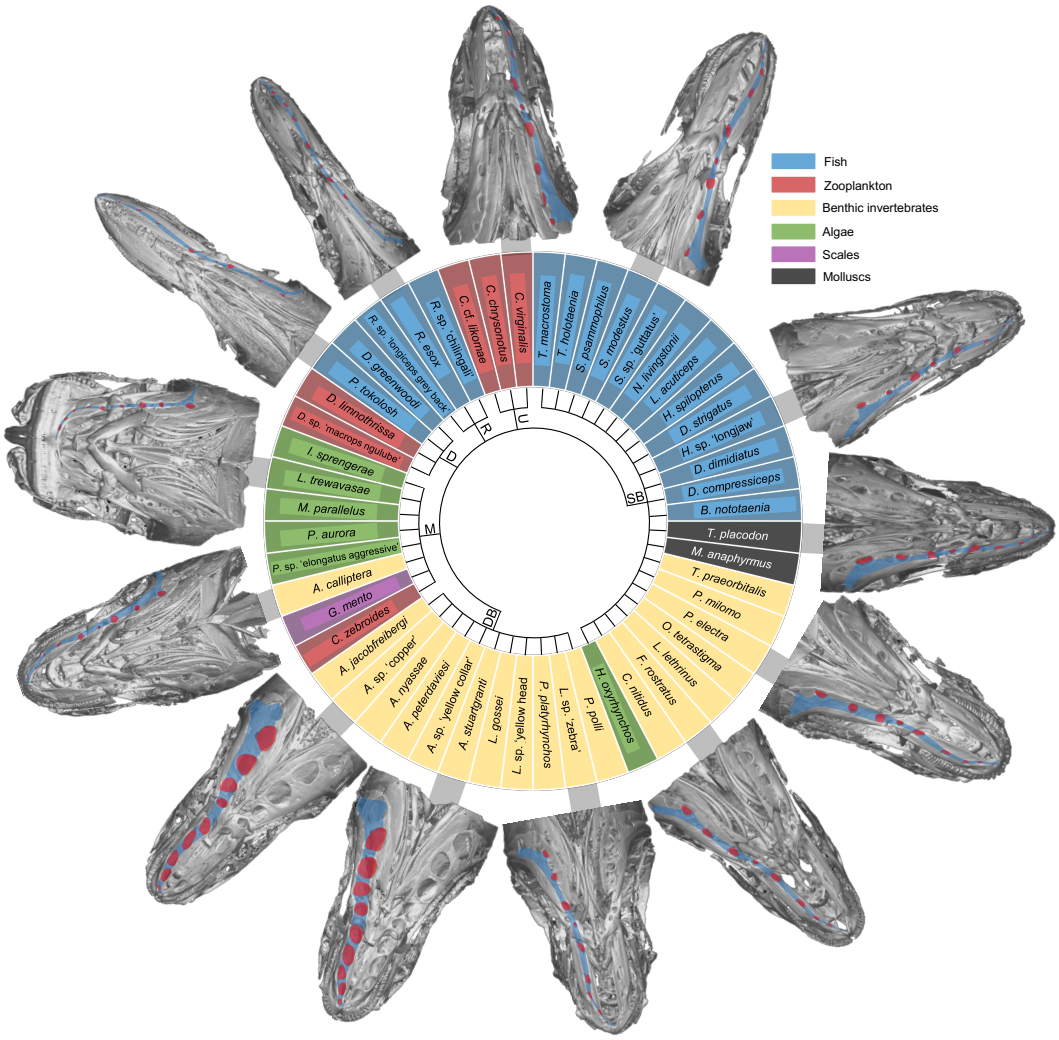


Figure 1.9 Variation in cranial canal lateral line system morphology among the cichlid species of Lake Malawi. Edgley & Genner (2019). See original paper for a full list of species and details. The 52 individual species shown here are grouped into the six major evolutionary lineages in the flock (Malinsky et al. 2018), and colours indicate diet.

1.5 Aims of the research presented in this thesis

Through the research presented in this thesis, I aim to investigate the role of morphological diversification in evolution and speciation in fishes, by focusing on the cichlids of Lake Malawi and its surrounding rivers and satellite lakes. I first aim to build upon the initial research on the diversification of the lateral line system within the Malawi flock (Edgley and Genner 2019). In Chapter 2, I focus on describing morphological diversity of the lateral line system in a population at the outset of ecological speciation and investigate its role in promoting diversification during initial population divergence. In Chapters 3 and 4 I focus on laboratory-bred segregating F_2 interspecific crosses of Lake Malawi species with divergent lateral line system morphologies, and using genomic and behavioural analyses respectively to reveal insights into the timing and dynamics of the process of morphological diversification, and its adaptive function. In Chapter 5

I broaden the work to study which traits are under selection during initial population divergence. I focus on microevolutionary dynamics of early-stage diversification in a small satellite lake, studying a range of phenotypic traits, including lateral line system phenotypes.

Sampling site at Lake Masoko, Tanzania, October 2019
Photo: Madeleine Carruthers

Chapter 2

Early lateral line system diversification during ecological speciation in cichlid fishes

2.1 Abstract

The mechanosensory lateral line system is an important sensory modality in fishes, informing many behaviours related to survival, including finding food and navigating in darker or more turbid environments. Although vision is known to play an important role in early stages of speciation of fishes, the role of mechanoreception via the lateral line has not been determined. Given its ecological importance, we may expect lateral line morphology to be under disruptive selection during ecological speciation. Here, we quantify the lateral line system morphology of two ecomorphs of the cichlid fish *Astatotilapia calliptera* in an isolated crater Lake. These ecomorphs have diverged from common ancestry within the past 1,000 years, and comprise part of the larger Lake Malawi haplochromine cichlid radiation that contains over 700 species. We show evidence of divergent lateral line system morphologies between these ecomorphs, suggesting disruptive selection on aspects of lateral line structures in the early stages of incipient speciation. Specifically, we find that the zooplanktivorous benthic ecomorph that dominates the deeper waters of the lake has large cranial canals, while the nearshore invertebrate-feeding littoral ecomorph that dominates the shallower waters of the lakes has narrow cranial canals. We find no evidence for divergence in either the number of superficial or trunk canal neuromasts. We show the magnitude of this variation in the context of the nearby Lake Malawi radiation of cichlids. These results provide the first evidence of a key role for this often-overlooked sensory modality at the early stages of the ecological speciation process in cichlid fishes, giving insights into its role in the process of adaptive radiation.

2.2 Author Contributions

Duncan Edgley microCT scanned, reconstructed and landmarked all specimens, imaged and analysed all neuromast data, conducted all statistical analyses and wrote the manuscript. This chapter was conceived, conducted and written under the supervision of **Martin Genner**.

Fish used for the CT scan dataset were caught by **M. Emília Santos**, **Gregoire Vernaz**, and **George Turner** in 2018. Whole genome shotgun sequencing for categorisation of subpopulations and genomic PCA was conducted by **Tyler Linderoth**, **Richard Durbin**, **Bettina Fischer**, and **Duncan Edgley**. CT scanning, reconstruction and landmarking and subsequent analyses were all conducted by **Duncan Edgley**. Fish used for neuromast imaging, body shape PCA and LPJ PCA were caught by **Duncan Edgley** and **Madeleine Carruthers** in October 2019, with assistance of **Joseph Masore**, **Nestory P. Gabagambi**, **Charles Malela**, **Thomas Katerso** and **Zacharia J. Mwampwani**. **Andy Saxon** digitised landmarks for gross body shape and LPJ shape (including microCT scanning and reconstruction). **Madeleine Carruthers** performed these PCA analyses. **Thomas Davies** and **Elizabeth Martin-Silverstone** assisted with microCT scanning. Lake Malawi fish for CT scan comparisons were collected by **Martin Genner**. microCT scanning, landmark digitisation and data analysis were by **Duncan Edgley**, outlined in Edgley & Genner (2019). Lake Malawi fish for neuromast comparisons were provided by **Domino Joyce**, **Alan Smith**, and **George Turner**. They were all imaged, processed and data was collected by **Duncan Edgley**. **Nestory P. Gabagambi**, **Asilatu Shechonge** and **Mary Kishe** helped with logistics and planning during fieldwork.

2.3 Introduction

To understand how organisms diversify and eventually become separate species, it is important to investigate how sensory systems adapt to maximise fitness within their environments (Seehausen et al. 2008). In principle, during ecological speciation, selection will tune sensory systems to the specific requirements of their respective niche, whether this is related to feeding, navigating, or finding a mate and reproducing (Rundle and Nosil, 2005), and there is considerable evidence for selection leading to modifications of sensory systems within the context of rapid speciation events (Seehausen 2006). Notably, the proliferation of forms characterising adaptive radiations is often accompanied by modifications to sensory systems (Carleton et al. 2005; Seehausen et al. 2008; Terai et al. 2017; Musilová et al. 2019), and can even extend to substantial changes in the morphology of these sensory organs themselves (Edgley and Genner, 2019).

Aquatic environments can be visually and hydrodynamically noisy, so fish may need to rely on their multiple integrated and adaptable sensory modalities to survive (Bleckmann and Zelick, 2009; Brumm, 2013; Mogdans, 2019; Attwell et al. 2021). Research on the evolution of sensory systems in fish has largely focused on vision, and this has clearly demonstrated that adaptation of visual systems is important for both initial divergence and the maintenance of reproductive isolation in sympatry (Carleton et al. 2005; Seehausen et al. 2008; Musilová et al. 2019). However, the ability of fishes to survive and adapt to their environment is also dependent on the detection of water flow through the mechanoreceptive lateral line system (Engelmann et al., 2000; Webb, 2014; Mogdans, 2019). Found in all fishes, this sensory modality is used for location and identification of prey, predators and conspecifics (Bleckmann et al. 1989; Janssen, 1996; Schwalbe et al. 2012; Schwalbe and Webb, 2014; Butler and Maruska, 2015; Edgley and Genner, 2019), as well as detecting substrate proximity and mediating rheotaxis (Montgomery et al. 1997; Suli et al. 2012; Yoshizawa et al. 2014). It is utilised in many fish behaviours, and can show substantial variation within rapidly speciating clades, such as the adaptive radiations of cichlid fishes of the African Great Lakes (Edgley and Genner, 2019), implying it is likely to be under disruptive selection during initial species divergence.

Moreover, due to the range of functions of the lateral line, we may expect selection to act on this trait across multiple axes of niche divergence during early stages of adaptive radiation, regardless of which selection pressure is most influential at this stage. For instance, if populations begin to diverge first through changes in in microhabitat use (Ronco et al., 2021), they are likely to require different mechanoreceptive capabilities to adapt to the variable depths, levels of turbidity and levels

of abiotic flow in their new environments (Mogdans, 2019). Likewise, if speciation is initially characterised by segregation into trophic niches, we would also expect accompanying disparity in lateral line morphology (Cooper et al. 2010). In Lake Malawi cichlids, for example, trophic resources range from being highly motile (fish), to cryptic (infaunal invertebrates) or sessile (algae) (Turner, 1996; Genner and Turner, 2005; Konings, 2016), and any signals generated be these food items will lead to different requirements of lateral line system structures for the species that consume them (Montgomery, 1989; Schwalbe et al., 2012).

There are multiple examples of the functional morphological diversification of the lateral line systems in fishes (Wark & Peichel 2010; Yoshizawa et al. 2014; Nickles et al. 2020). Variation in the morphology of the lateral line system has also been described across the major clades of haplochromine cichlids in Lake Malawi, suggesting that it may have a role in enabling adaptive radiation (Edgley and Genner, 2019). The expanded canal lateral line system of the Lake Malawi cichlid *Aulonocara stuartgranti* has been shown to facilitate feeding in the dark, implying that it is at least partially a dietary adaptation (Schwalbe et al. 2012; Schwalbe and Webb, 2014; Edgley and Genner, 2019). In other cichlids, the lateral line system has been shown to mediate interactions between males during aggressive territorial interactions (Butler and Maruska, 2015). Although the system's morphological and functional diversity in cichlids implies an important role in adaptive diversification, there is little evidence of adaptive lateral line system divergence in a microevolutionary context. Specifically, the ecological and behavioural correlates of lateral line variation are not fully resolved, and it is not known whether diversification takes place early or later in the species process.

Here, we test whether diversification in lateral line system morphology takes place during the early stages of ecologically-associated cichlid fish speciation. We focus on a small crater lake in Tanzania, Lake Masoko (also known as Lake Kisiba) (Figure 2.1A), which hosts two genetically divergent 'ecomorphs' of *Astatotilapia calliptera* (Malinsky et al. 2015). These two ecomorphs have diversified into their respective forms within the last 1000 years, following colonisation of the lake by riverine ancestors within the past 10,000 years (Barker et al. 2003) (Figure 2.1A & 2.1D). They exhibit distinct depth preferences and inhabit ecologically different environments (Figure 2.1B-C). The "littoral" ecomorph dominates the shallow brightly lit littoral habitat (<10m), whereas the "benthic" ecomorph dominates the deeper dimly lit benthic habitat (>20m) (Figure 2.1B-C). In addition, the two ecomorphs have distinct diets (Figure 2.1E), male nuptial colouration (Figure 2.1D), body shapes (Malinsky et al. 2015; Carruthers et al. 2022) (Figure 2.1G) and lower

pharyngeal jaw morphology (Malinsky et al. 2015; Carruthers et al. 2022; Vernaz et al. 2022) (Figure 2.1G).

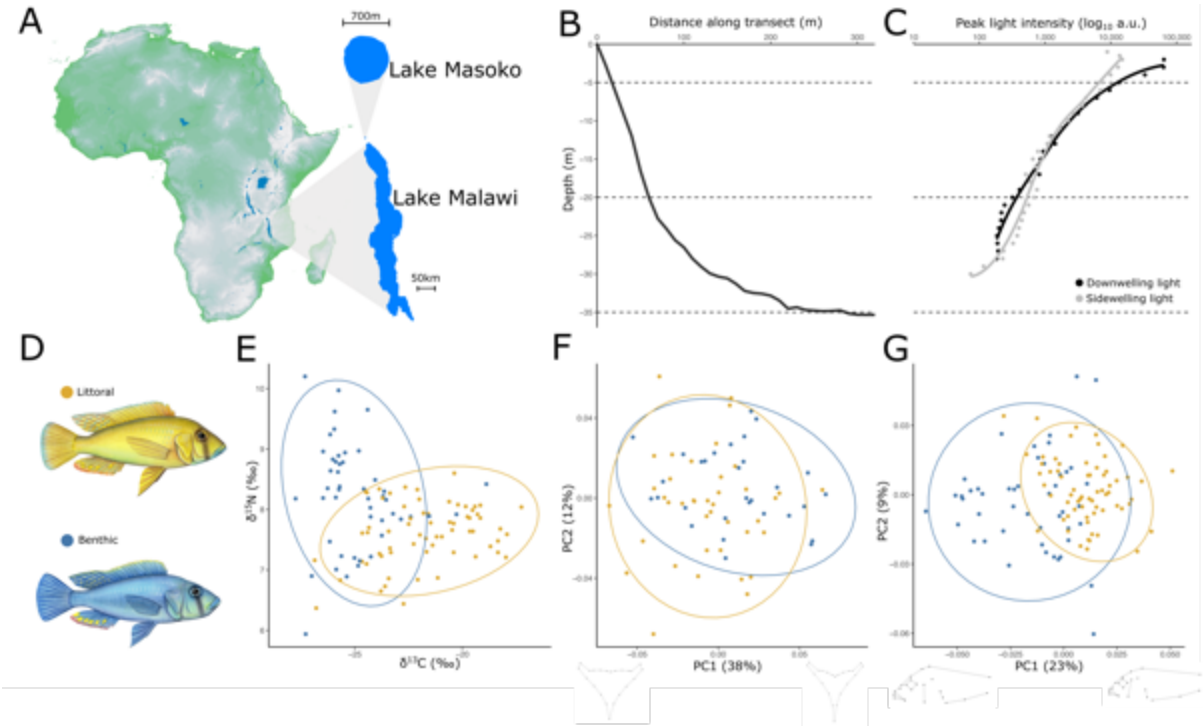


Figure 2.1 An overview of the Lake Masoko *Astatotilapia calliptera* system. A) The location of Lake Masoko within Africa, relative to Lake Malawi. (B) Bathymetric profile of Lake Masoko, averaged across four transects (data from Turner et al., 2019; data collected in 2018). (E) The peak light intensity by depth for sidewelling light (grey) and downwelling light (black) (data collection in summer 2018). (D) Illustrations of males of the two ecomorphs found in Lake Masoko, the yellow littoral ecomorph (above) that dominates the habitat <10m, and the blue benthic ecomorph (below) that dominates the habitat >20m. Paintings are by Julie Johnson (https://www.lifesciencestudios.com/). (E) Stable isotope analysis for ^{13}C and ^{15}N in Lake Masoko cichlids (n=113). (F) Principal component analysis of LPJ shape in Lake Masoko cichlids (n=70). (G) Principal Component Analysis of body shape in Lake Masoko cichlids (n=113). For F and G also shown are 95% confidence ellipses for benthic and littoral ecomorphs, and shape changes along the first PC axis. Data is from Carruthers et al. 2022.

The Lake Masoko *A. calliptera* system provides the opportunity to study the dynamics of the early species of speciation within a geographically constrained and but well characterised environment, providing insight into how speciation takes place within larger-scale adaptive radiations. We have good evidence to suggest that the Lake Masoko ecomorphs have diversified in visual system as an adaptation to the light levels in their respective environments (Sugawara et al. 2005; Malinsky et al. 2015), but whether they are being exposed to divergent selection on their lateral line systems is not known. Hence, in this study we comprehensively investigate variation in both the canal and superficial components of the lateral line system. We then place these results in context by comparing lateral line morphology of the two Lake Masoko ecomorphs with the variation in ecologically-divergent groups of species from the older (1Ma) and larger (700+ species) Lake Malawi radiation (of which *A. calliptera* is a member) (Genner and Turner, 2011; Malinsky et al.

2018). This enables specific insight into whether patterns of early divergence we observe in Lake Masoko are reflective of those we see in the larger radiation.

2.4 Materials and Methods

2.4.1 Sample collection

For the visualisation and imaging of superficial and canal neuromasts, 75 specimens of the cichlid *Astatotilapia calliptera* were caught at Lake Masoko (9°20'00.4"S 33°45'19.6"E) in October 2019 (Figure 2.1A) using SCUBA, and depth gauges to determine capture depth. 53 "shallow" individuals were caught at a target range of <5m, and 22 "deep" individuals were caught at >20m (Figure 2.1D; Figure 2.1F). Littoral fish were immediately transported to the Tanzania Fisheries Research Institute (TAFIRI) laboratory in Kyela, Tanzania in aerated barrels, whereas benthic fish were depressurised in holding barrels over two days before transportation. Fish were kept in aerated holding tanks at TAFIRI Kyela before being processed (outlined below under "Neuromast imaging").

For the computed tomorgraphy (CT) scanning analysis, 199 individuals were caught at Lake Masoko in August 2018. Of these 51 were caught in the shallows (<5m), 83 were from the intermediate (midwater) depths (5-20m), and 65 were caught in deep waters (>20m). Analyses were conducted on fish grouped by the depth they were collected, but also based on genetically-defined subpopulations. These were determined using whole genome shotgun sequencing data, with reads aligned to the fAstCal1.2 reference genome (accession GCF_900246225.1). Following sequencing using Illumina HiSeq, variant calling and quality filtering, a genomic principal components analysis (PCA) (Munby et al., 2021) revealed clear separation of ecomorphs by the primary axis of genetic variation: PC1 (2.158%) (Figure S2.1). Following Munby et al. (2022) we use PC1 of this genomic PCA to define our genomic subpopulations. "Benthic" individuals are defined as PC1 > 0.04 (n = 64). "Littoral" individuals are defined as PC1 < -0.02 (n = 115); "intermediate" individuals are defined as PC1 between -0.02 and 0.04) (n=20). For all subsequent analyses, "littoral", "intermediate" and "benthic" refer to the genomic subpopulations, whereas "shallow", "midwater" and "deep" refer to capture depth.

2.4.2 Cranial canal pore morphology

To visualise the anatomy of the cephalic lateral line system (defined as anterior of the preoperculum), the heads of all individuals were microCT scanned using a Nikon XTH225ST micro-computed tomography (micro-CT) system at the University of Bristol. Each scan covered two individuals and used 3,141 projections. Voxel size for each scan was 20-30 μm. Scan resolution was determined by the size of the cephalic lateral line canal pores, which were in turn determined by specimen size. Preliminary scans using similar specimens, and experience from previous scanning (Edgley and Genner, 2019) determined an appropriate resolution for subsequent scans. Image stacks were imported into VGStudio MAX 3.3.6 (Volume Graphics GmbH, 2016) and reconstructed into a 3D model. 2D images of the reconstructed model were captured from the ventral head perspective (showing the pores of the mandibular, infraorbital canals, and the lower arm of the preopercular canal), and the lateral head perspective on the left side (showing the pores of the lateral arm of the preopercular canal, and the otic canal) (Figure 2.2A-B).

A landmark-based morphometric approach was used to estimate the area of the cranial canal pores of each specimen. Using tpsDig 2.31 (Rohlf 2015), curves of sliding semilandmarks were drawn around the circumference of each pore, anchored by landmarks at its anterior limit. These curves were resampled resulting in ten equidistant semilandmarks for each pore. All landmarks were digitised by the same individual with only short breaks between landmarking sessions to minimise human error. We tested for digitisation error by redigitising 20 individuals' cranial canal pores, and testing for differences between our dataset and the re-digitised scans. We found no significant difference between the two. After conversion from semilandmarks to landmarks in tpsUtil32 (Rohlf 2015), landmark coordinates were imported into R 3.1.3 (R Core Team, 2020) using the package 'geomorph' (Adams et al. 2022). Image scale was accounted for when importing landmarks. The tenth landmark for each pore was discarded as its coordinates are the same as the first. The remaining nine landmarks were treated as vertices of a polygon, of which the area was calculated using the polyarea function in the 'geometry' package (Roussel et al. 2022). Resulting areas are approximations of the true pore area. Although being a slight underestimate, we were consistent in our methodology across all specimens. We used this method to estimate the size of the five anterior-most pores of the mandibular canal (Figure 2.2E), the seven pores of the preopercular canal (Figure 2.2C & 2.2F), the six pores within the infra-orbital canal (Figure 2.2D), and two pores of the otic canal (Figure 2.2C).

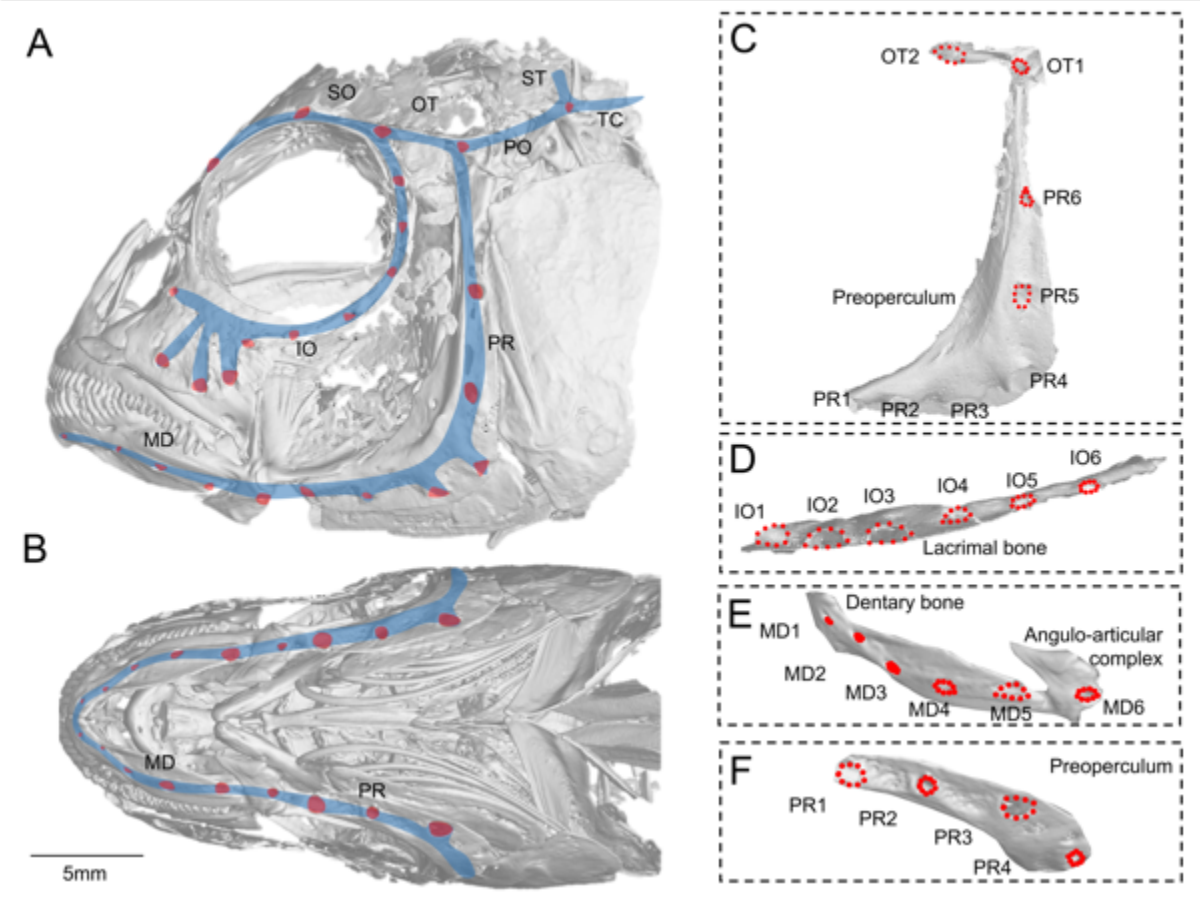


Figure 2.2 The cranial lateral line canal system of *Astatotilapia calliptera* from Lake Masoko. (A) Annotated microCT scan of the lateral head, showing the cranial canal pores (red) and the approximate location of the cranial canals (blue). MD = mandibular canal; PR = preopercular canal; IO = infraorbital canal; SO = supraorbital canal; OT = otic canal; PO = post-otic canal; ST = supratemporal canal; TC = anterior trunk canal. (B) Annotated microCT scan of the ventral view of the head. (C-F) The landmarking regime of this study, used to approximate canal pore area. Red points are the landmark locations used; each pore is bounded by nine fixed landmarks placed equidistant along the outer edge of the pore. (C) The isolated preoperculum and otic canals, and labelled pores. (D) the isolated infraorbital canal and pores. (E) The isolated mandibular canal and pores, located within the dentary bone and angulo-articular bones. (F) The ventral-facing pores of the preopercular canal found within the preoperculum.

2.4.3 Neuromast imaging

For imaging the neuromasts of our 75 fish, fish were stained in 0.008% solution of the fluorescent dye DASPEI [2-(4-(dimethylamino)styryl)-N-ethylpridinium iodide; Fisher Scientific]. DASPEI is a vital mitochondrial dye, commonly used for the staining of epidermal mechanoreceptors and electroreceptors in teleosts and larval amphibians (Bereiter-Hahn 1976; Butler et al. 2016). It is commonly used to visualise both superficial and canal neuromasts in many fish groups, including cichlid fishes (Webb et al. 2014; Butler and Maruska, 2015; Butler et al. 2016). After submersion in DASPEI solution, fish were subsequently euthanised using MS-222 (an approved Schedule 1

method). Neuromasts of each fish were imaged on the left side of the body using a Canon EOS 500D DSLR camera and a Sigma 18-200 mm f/3.5-6.3 lens. Photos were taken in dark conditions with a royal blue lamp (465 nm) under a yellow glass longpass filter (500 nm) to remove interference. Images for each individual fish were stitched together in fiji 1.51 (Schindelin et al. 2012) and ImageJ (Schneider et al. 2012) (Figure 2.3C). Superficial and canal neuromasts were counted visually for each specimen where they were clearly visible (Figure 2.3B). For all analyses "Total head superficial neuromasts" is defined as the sum of neuromasts in the forehead, nose, pre-gill, post-gill and lower jaw regions (Figure 2.3B). "Total trunk superficial neuromasts" defined as the sum of the superficial neuromasts in the anterior and posterior trunk canals (Figure 2.3A-B) and includes any superficial neuromasts in the trunk region but not in close proximity to trunk canal neuromasts. "Total trunk canal neuromasts" is calculated as the sum of canal neuromasts in both the anterior and posterior trunk canals (Figure 2.3B).

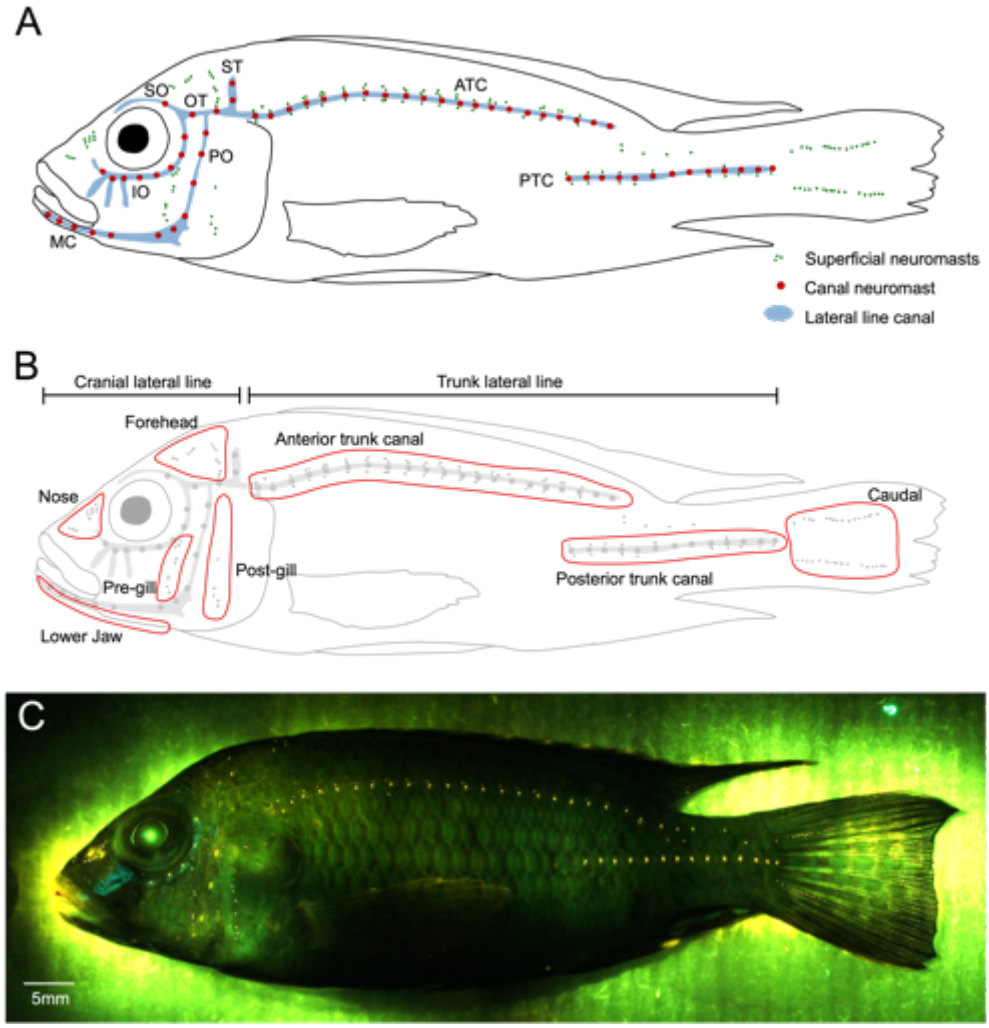


Figure 2.3 Neuromast patterning of *Astatotilapia calliptera* from Lake Masoko. (A) Schematic overview of the lateral line system of a deep-caught individual (MDP006 in this dataset). Locations of superficial (green) and canal neuromasts (red) are shown, along with approximate location of lateral line canals (blue). MC = mandibular canal; IO = infraorbital canal; PO = preopercular canal; SO = supraotic canal; OT = otic canal; ST = supratemporal canal; ATC = anterior trunk canal; PTC =

posterior trunk canal. (B) The neuromast counts used in this analysis. For both the anterior and posterior trunk canal, both superficial neuromasts and canal neuromasts were counted. For the cranial lateral line system, only superficial neuromasts were counted. (C) A stitched photograph of a DASPEI stained *Astatotilapia calliptera*. Note: photographs were also taken of the ventral view of the lower jaw, as the neuromasts there are not visible from this perspective.

2.4.4 Comparisons with Lake Malawi cichlids

In addition to *Astatotilapia calliptera* from Lake Masoko, we also analysed the pore areas from the four focal cranial canals (mandibular, preopercular, infraorbital and otic) in 52 species from Lake Malawi (n=1 for each species), which were representative of the major ecomorphologically divergent clades in the species flock: the shallow rocky shore-dwelling mbuna; the shallow open water zooplanktivorous utaka; the deep water predatory *Diplotaxodon*; the open water *Rhamphochromis*; the shallow water benthic group; and the deep water benthic group (Malinsky et al. 2018). The microCT scans used for this analysis were taken from Edgley & Genner (2019). microCT scanning, landmark digitisation and subsequent analysis were conducted as described above.

For neuromast comparisons we included several species representative of the diversity found in the major clades present in Lake Malawi (Konings, 2016; Malinsky et al. 2018). We imaged 12 different species within 6 groups, summarised in Table 2.2. Individuals were sourced from or bred from existing stock in aquaria in the UK, at either Bangor University, the University of Hull or the University of Bristol. Between four and ten individuals of each of these were imaged in the same manner as Lake Masoko *A. calliptera* as outlined above. Trunk neuromasts were counted using the same procedures as outlined above (Figure 2.3).

Table 2.1 Species included in our CT scanning dataset. All individuals were wild caught in either Lake Malawi or Lake Masoko. *Wild caught by S.M. Grant, location unknown. **Putatively undescribed species.

Group	Species	Sample size	Collection year	Sampling Location
Littoral	*Astatotilapia calliptera* (Masoko littoral)	115	2018	Lake Masoko
Intermediate	*Astatotilapia calliptera* (Masoko intermediate)	20	2018	Lake Masoko
Benthic	*Astatotilapia calliptera* (Masoko benthic)	64	2018	Lake Masoko
mbuna	*Astatotilapia calliptera*	1	2011	Lake Itamba, Tanzania
	Cynotilapia zebroides	1	1997	Monkey Bay, Malawi
	Genyochromis mento	1	1997	Nkhata Bay, Malawi
	Iodotropheus sprengerae	1	2011	Chiofu, Malawi
	Labeotropheus sp. 'variable eastern'	1	2012	Unknown*
	Melanochromis loriae	1	1997	Nkhata Bay, Malawi
	Pseudotropheus aurora	1	2014	Cape Maclear, Malawi
	Pseudotropheus sp. 'elongatus aggressive'	1	2014	Cape Maclear, Malawi
utaka	*Copadichromis chrysonotus*	1	2014	Mangochi, Malawi
	Copadichromis likomae	1	2014	Cape Maclear, Malawi
	Copadichromis mloto	1	2014	Cape Maclear, Malawi
Rhamphochromis	*Rhamphochromis esox*	1	2004	Nkhata Bay, Malawi
	Rhamphochromis sp. 'chilingali'	1	2004	Lake Chilingali, Malawi
	Rhamphochromis sp. 'longiceps grey-back'	1	2005	Dwangwa, Malawi
Diplotaxodon	*Diplotaxodon greenwoodi*	1	2004	Cape Maclear, Malawi
	Diplotaxodon limnothrissa	1	2004	Cape Maclear, Malawi
	Diplotaxodon sp. 'macrops ngulube'	1	2005	Nkhata Bay, Malawi
	Pallidochromis tokolosh	1	2005	Kasuza, Malawi
Shallow Benthic	*Buccochromis nototaenia*	1	2014	Lake Malombe, Malawi
	Ctenopharynx nitidus	1	2014	Cape Maclear, Malawi
	Dimidiochromis compressiceps	1	2011	Mangochi, Malawi
	Dimidiochromis dimidiatus	1	2014	Cape Maclear, Malawi
	Dimidiochromis strigatus	1	2014	Mangochi, Malawi
	Fossarochromis rostratus	1	1997	Nkhata Bay, Malawi
	Hemitaeniochromis sp. 'longjaw'**	1	2014	Mangochi, Malawi
	Hemitaeniochromis spilopterus	1	2005	Monkey Bay, Malawi
	Hemitilapia oxyrhynchos	1	2014	Mangochi, Malawi
	Lethrinops lethrinus	1	2014	Lake Malombe, Malawi
	Lichnochromis acuticeps	1	2014	Cape Maclear, Malawi
	Mylochromis anaphyrmus	1	2014	Cape Maclear, Malawi
	Nimbochromis livingstonii	1	2012	Unknown*
	Otopharynx tetrastigma	1	2014	Cape Maclear, Malawi
	Placidochromis electra	1	2012	Unknown*
	Placidochromis milomo	1	2005	Monkey Bay, Malawi
	Sciaenochromis psammophilus	1	2004	Maldeco Fisheries,
	Stigmatochromis modestus	1	2014	Cape Maclear, Malawi
	Stigmatochromis sp. 'guttatus'	1	1996	Nkhata Bay, Malawi
	Taeniochromis holotaenia	1	2014	Cape Maclear, Malawi
	Taeniolethrinops praeorbitalis	1	2014	Mangochi, Malawi
	Trematocranus placodon	1	2005	Monkey Bay, Malawi
	Tyrannochromis macrostoma	1	1997	Nkhata Bay, Malawi
Deep Benthic	*Alticorpus peterdaviesi*	1	2005	Monkey Bay, Malawi
	Aulonocara jacobfreibergi	1	2014	Cape Maclear, Malawi
	Aulonocara nyassae	1	2014	Cape Maclear, Malawi
	Aulonocara sp. 'copper'	1	2005	Monkey Bay, Malawi
	Aulonocara sp. 'yellow collar'	1	2014	Cape Maclear, Malawi
	Aulonocara stuartgranti	1	2012	Unknown*
	Lethrinops gossei	1	2005	Tukombo, Malawi
	Lethrinops sp. 'yellow head'	1	2014	Cape Maclear, Malawi
	Lethrinops sp. 'zebra'	1	2005	Nkhata Bay, Malawi
	Placidochromis platyrhynchos	1	2014	Cape Maclear, Malawi
	Placidochromis polli	1	2005	Monkey Bay, Malawi

Table 2.2 Neuromast imaging specimens used in this study.

Group	Species	Sample size	Sourced from	Imaging location
Shallow	*Astatotilapia calliptera* (Masoko Shallow)	53	Lake Masoko (wild)	Kyela, Tanzania
Deep benthic	*Astatotilapia calliptera* (Masoko Deep)	22	Lake Masoko (wild)	Kyela, Tanzania
A. calliptera "Salima"	*Astatotilapia calliptera* (Salima)	6	Lake Malawi (aquarium)	Bangor University
mbuna	*Maylandia zebra*	6	Lake Malawi (aquarium)	Bangor University
	Maylandia estherae	5	Lake Malawi (aquarium)	Hull University
	Labeotropheus trewavasae	5	Lake Malawi (aquarium)	Hull University
utaka	*Nyassochromis microcephalus*	5	Lake Malawi (aquarium)	Hull University
Rhamphochromis	*Rhamphochromis* sp. "chilingali"	5	Lake Malawi (aquarium)	Bangor University
	Rhamphochromis longiceps	5	Lake Malawi (aquarium)	Hull University
Diplotaxodon	*Diplotaxodon limnothrissa*	5	Lake Malawi (aquarium)	Hull University
Shallow benthic	*Nimbochromis livingstonii*	5	Lake Malawi (aquarium)	Hull University
	Protomelas sp. "johnstoni solo"	5	Lake Malawi (aquarium)	Hull University
	Otopharynx lithobates	10	Lake Malawi (aquarium)	Bristol University
Deep benthic	*Aulonocara stuartgranti*	13	Lake Malawi (aquarium)	Bangor University

2.4.5 Statistical analyses

To test for significant differences between ecomorphs or capture depths, we used general linear models (GLMs) for each morphological variable. For each of the four cranial canals (mandibular, preopercular, infraorbital and otic), we calculated the mean pore area. Within each canal, we used $\log_{10}$ transformed mean pore area as the response variable and subpopulation (littoral, benthic or intermediate) and sex as predictor variables for gaussian GLMs. In addition, we created models for each including capture depth as the predictor variable (shallow < 5m, midwater 5-20m, deep > 20m). For each linear model we used Tukey's HSD (honestly significant difference) *post-hoc* tests, using Bonferroni correction for multiple testing. For all models, $\log_{10}$ standard length and sex were included as covariates to account for their effects. Plots of pore areas are partial residuals from these models, showing differences according to subpopulation or depth of capture, while accounting for standard length and sex. We similarly tested for significant differences in neuromast counts between capture depths using a generalized linear model (GLM) with a poisson fit, including capture depth (shallow < 5m; deep > 20m), $\log_{10}$ standard length and sex as predictor variables.

2.5 Results

2.5.1 Cranial canal pore morphology

There were significant differences between genetic subpopulations in the mean pore area of the mandibular canal in all four measured traits when accounting for variation in standard length, with a consistent pattern of significantly larger pore areas in the benthic fish than the littoral fish (Table 2.3; Figure 2.4). Genetically intermediate fish had significantly larger pore areas than the benthic fish in only preopercular and infraorbital canal pore size (Figure 2.4C & 2.4E), while genetically benthic fish had significantly larger pore areas than intermediate fish in preopercular, infraorbital and otic canal pore size (Figure 2.4C, 2.4E & 2.4G). Results were largely similar when grouping individuals by capture depth when accounting for variation in standard length (Table 2.4). In all four measured traits, fish caught deep (>20m) had significantly larger pores than those caught in the shallow water (<5m) (Table 2.4; Figure 2.4). Fish from the midwater zone had significantly larger pores than shallow waters in infraorbital and otic canal pore size. Fish from the benthic zone had significantly larger pores than shallow waters in all four canals we measured. We found no evidence of consistent variation in pore sizes between the sexes (Tables 2.3 & 2.4).

Table 2.3 Models of mean pore area measurements for four cranial canals, testing for differences between genetically characterised subpopulations. L=littoral; I=intermediate; B=benthic; ***p<0.001; **p<0.01; *p<0.05

Response	Predictors	Sum of squares	F	df	residual df	P value		Tukey contrasts	Post-hoc p-value
Log mean mandibular pore area	Subpopulation	0.733	48.96	2	194	**< 0.001*** **		L-B	**< 0.001*** **
	log_{10}(standard length)	1.810	241.77	1	194	**< 0.001*** **		I-B	1.000
	Sex	0.004	0.489	1	194	0.485		I-L	0.138
Log mean preopercular pore area	Subpopulation	1.346	107.20	2	194	**< 0.001*** **		L-B	**< 0.001*** **
	log_{10}(standard length)	1.683	268.17	1	194	**< 0.001*** **		I-B	**0.010****
	Sex	0.001	0.156	1	194	0.693		I-L	**0.024***
Log mean infraorbital pore area	Subpopulation	4.209	234.98	2	191	**< 0.001*** **		L-B	**< 0.001*** **
	log_{10}(standard length)	1.844	205.88	1	191	**< 0.001*** **		I-B	**< 0.001*** **
	Sex	0.000	0.039	1	191	0.844		I-L	**0.003****
Log mean otic pore area	Subpopulation	1.2380	70.36	2	194	**< 0.001*** **		L-B	**< 0.001*** **
	log_{10}(standard length)	1.7438	198.22	1	194	**< 0.001*** **		I-B	**0.007****
	Sex	0.007	0.757	1	194	0.385		I-L	0.785

Table 2.4 Models of mean pore area measurements for four cranial canals, testing for differences between fish from different capture depths. S=shallow; M=midwater; D=deep; ***p<0.001; **p<0.01; *p<0.05

Response	Predictors	Sum of squares	F	df	residual df	P		Tukey contrasts	Post-hoc p-value
Log mean mandibular pore area	Capture depth	0.805	53.63	2	194	**< 0.001*** **		S-D	**0.016***
	log_{10}(standard length)	1.733	231.03	1	194	**< 0.001*** **		M-D	**< 0.001*** **
	Sex	0.007	0.89	1	194	0.341		M-S	0.505
Log mean preopercular pore area	Capture depth	1.357	105.75	2	194	**< 0.001*** **		S-D	**< 0.001*** **
	log_{10}(standard length)	1.644	256.10	1	194	**< 0.001*** **		M-D	**< 0.001*** **
	Sex	0.001	0.156	1	194	0.694		M-S	1.000
Log mean infraorbital pore area	Capture depth	3.934	176.35	2	194	**< 0.001*** **		S-D	**< 0.001*** **
	log_{10}(standard length)	1.812	162.46	1	194	**< 0.001*** **		M-D	**< 0.001*** **
	Sex	0.063	5.63	1	194	**0.0186***		M-S	**< 0.001*** **
Log mean otic pore area	Capture depth	1.289	76.25	2	194	**< 0.001*** **		S-D	**< 0.001*** **
	log_{10}(standard length)	1.740	205.86	1	194	**< 0.001*** **		M-D	**< 0.001*** **
	Sex	0.028	3.25	1	194	0.073		M-S	**0.0432***

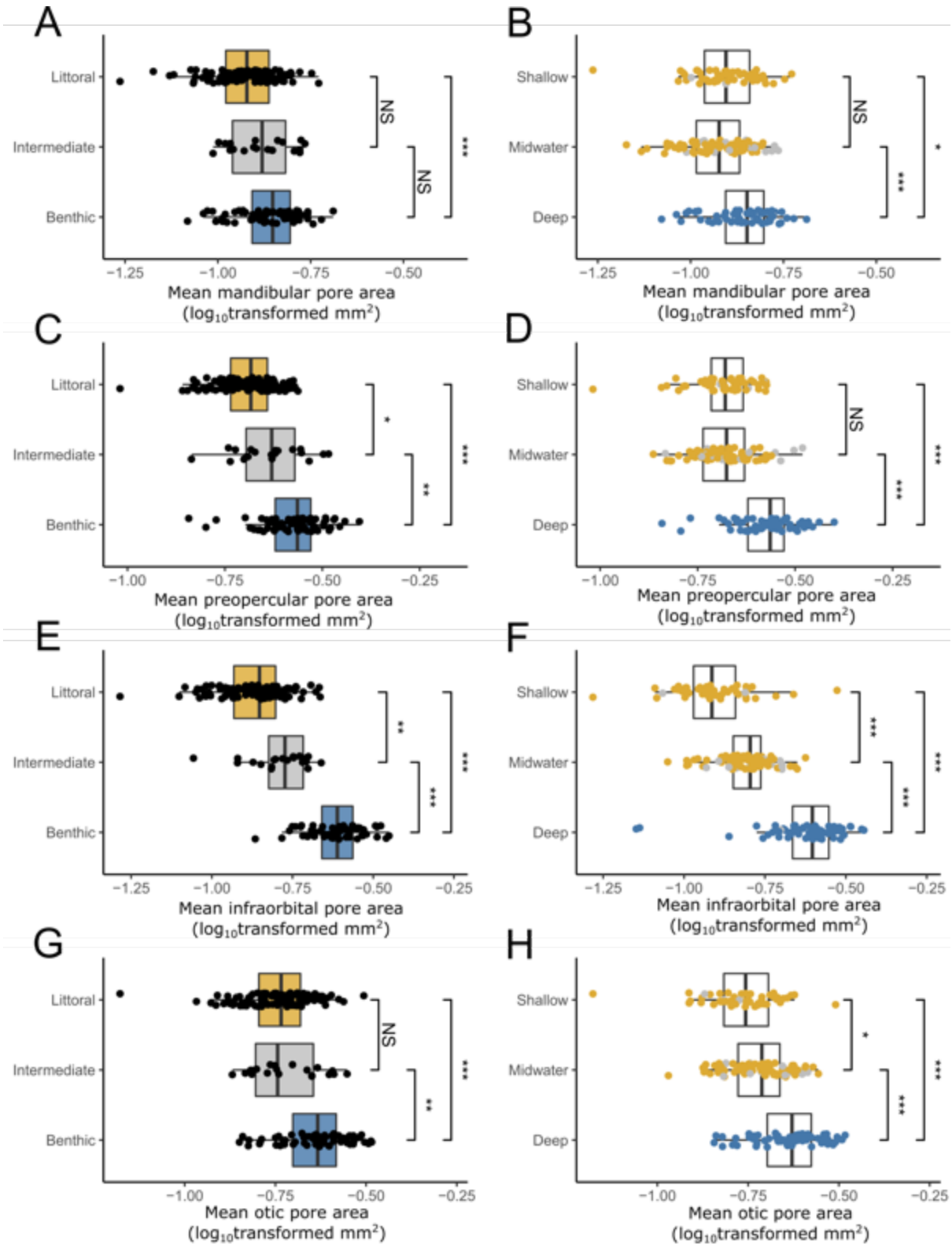


Figure 2.4 Mean pore sizes of cranial canals of the littoral, intermediate and benthic subpopulations of *Astatotilapia calliptera* from Lake Masoko. All plots show partial residuals of the response variable after accounting for standard length and sex. (A) Mean mandibular canal pore area grouped by genomic subpopulation. (B) Mean mandibular canal pore area grouped by capture depth. (C) Mean preopercular canal pore area grouped by genomic subpopulation. (D) Mean preopercular canal pore area grouped by capture depth. (E) Mean infraorbital canal pore area grouped by genomic subpopulation. (F) Mean infraorbital canal

pore area grouped by capture depth. (G) Mean otic canal pore area grouped by genomic subpopulation. (H) Mean otic canal pore area grouped by capture depth on. NS=not significant; *p>0.05; **p>0.01; ***p>0.001

2.5.2 Superficial and canal neuromasts

We found no significant differences between shallow and deep-caught fish in the number of superficial neuromasts on the head (Figure 2.5A), or the number of superficial neuromasts on the trunk (Figure 2.5B; Figure 2.2B). There was no difference between ecomorphs in the number of trunk neuromasts within both branches of the trunk canal (Figure 2.5C). When comparing the average number of superficial neuromasts found clustered around each canal neuromast of the trunk canal (both anterior and posterior arms), we found no significant differences between the two (Figure 2.5D). In addition, we found no evidence for significant differences in neuromast number by standard length or sex.

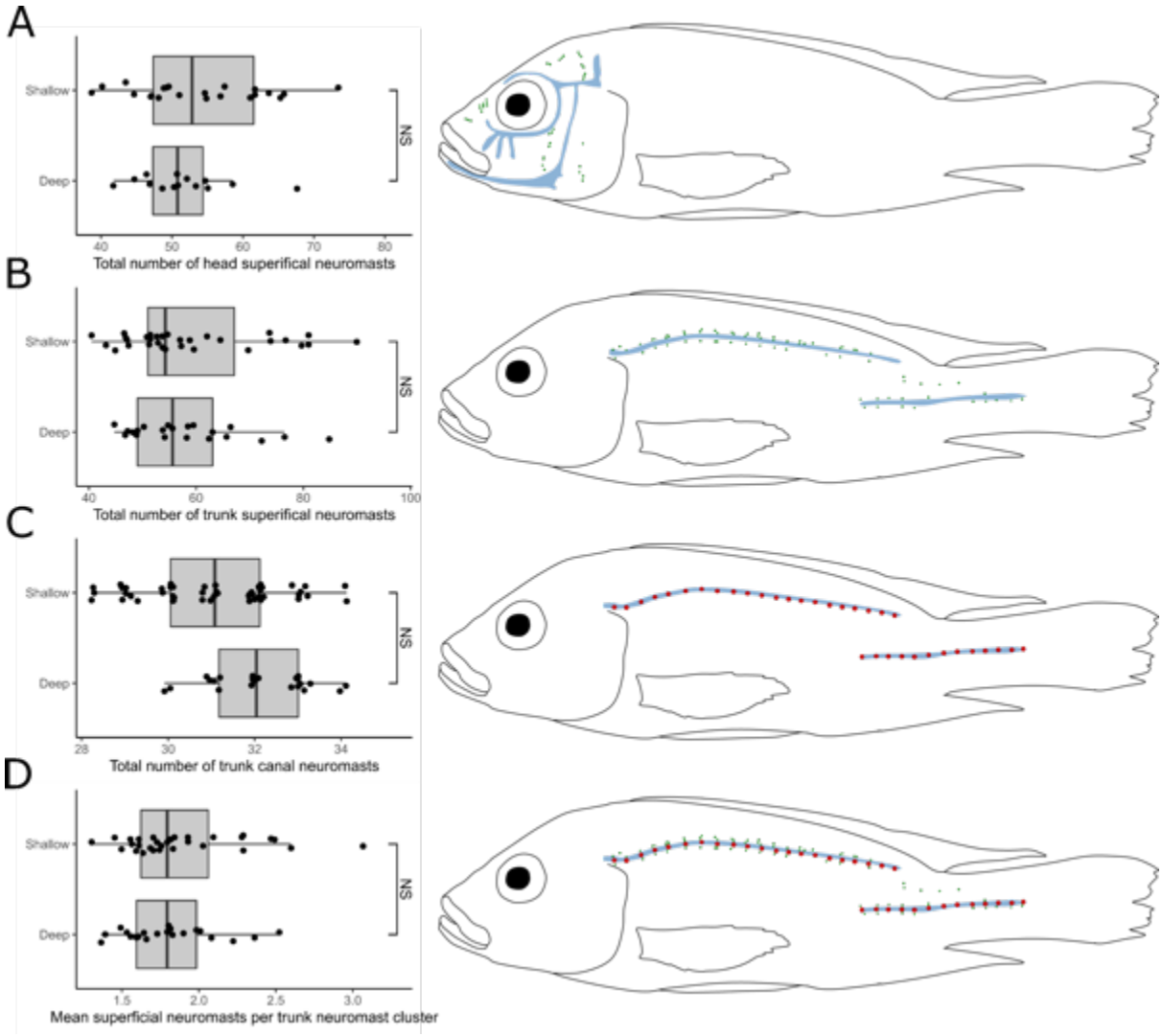


Figure 2.5 Neuromast counts for wild-caught *Astatotilapia calliptera* from Lake Masoko. Shallow individuals were caught above 5m, and Deep individuals were caught below 20m. (A) The total number of head superficial neuromasts. (B) The total

number of trunk superficial neuromasts. (C) The total number of trunk canal neuromasts. (D) The average number of trunk superficial neuromasts per canal neuromast for both the anterior and posterior trunk canals.

Table 2.5 Models of neuromast counts for superficial neuromasts and canal neuromasts of the trunk, testing for differences between shallow (<5m) and deep caught (>20m) individuals.

Response	Predictor	Estimate	Std. error	p-value
Total number of head superficial neuromasts	Capture depth	0.040	0.049	0.410
	$\log_{10}$(standard length)	-0.265	0.306	0.387
	Sex	-0.116	0.089	0.194
Total number of trunk superficial neuromasts	Capture depth	0.012	0.038	0.760
	$\log_{10}$(standard length)	-0.307	0.227	0.176
	Sex	0.052	0.054	0.330
Total number of trunk canal neuromasts	Capture depth	-0.038	0.046	0.417
	$\log_{10}$(standard length)	0.057	0.263	0.828
	Sex	-0.009	0.069	0.890
Mean number of superficial neuromasts per trunk cluster	Capture depth	0.048	0.213	0.820
	$\log_{10}$(standard length)	-0.359	1.271	0.777
	Sex	0.062	0.301	0.838

2.5.3 Variation in Lake Malawi

Evidence from Lake Malawi cichlids shows patterns of cranial canal morphology can be clade-specific (Edgley and Genner, 2019). To expand on this, and contextualise the extent of divergence in Lake Masoko, we calculated mean pore areas and counted neuromasts in representatives of the six major clades of Lake Malawi haplochromine cichlid radiation (Malinsky et al., 2018). Lineages which live more pelagic lifestyles, living and feeding in the open water (*Rhamphochromis*, utaka, and *Diplotaxodon*), tended to have relatively small canals with narrower openings to their environment (Figure 2.6A-D), but more neuromasts on the head and body (Figure 2.6E-H). *Rhamphochromis* and *Diplotaxodon* are primarily piscivorous predators, and utaka feed on zooplankton in the water column (Turner, 1996; Konings, 2016; Edgley and Genner, 2019). The broad "shallow benthic" and "deep benthic" clades, which are both characterised by associations with the lake sediment, include some individuals with larger cranial canals and pores, perhaps as an adaptation for feeding on prey buried on the sediment (Schwalbe et al., 2012; Konings, 2016; Edgley and Genner, 2019) (Figure 2.6A-D). Fish in these lineages tend to also have greater numbers of superficial and canal neuromasts (Figure 2.6E-H). The rocky-shore dwelling mbuna, including the generalist *A. calliptera*, tend to have narrower cranial canals with smaller pores, and fewer neuromasts (Figure 2.6).

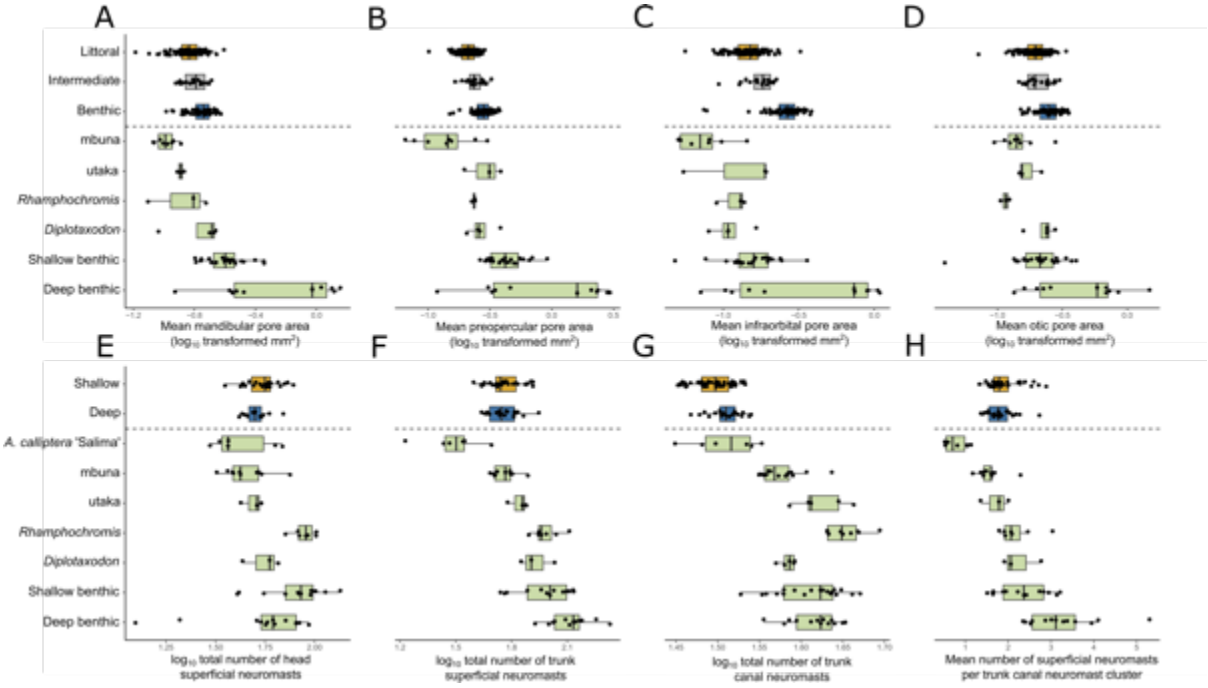


Figure 2.6 Comparisons of pore area and neuromast count measurements between Lake Masoko *Astatotilapia calliptera* and Lake Malawi individuals. Each point is an individual specimen, and all data are partialized residuals from statistical models, accounting for $\log_{10}$ standard length and sex (models outlined in table S2.1). (A) Mean mandibular pore area. (B) Mean preopercular pore area. (C) Mean infraorbital pore area. (D) Mean otic pore area. (E) Total number of head superficial neuromasts. (F) Total number of trunk superficial neuromasts. (G) Total number of trunk canal neuromasts. (H) Mean number of superficial neuromasts per trunk canal neuromast cluster.

2.6 Discussion

This study investigated the evidence for divergence in two components of the lateral line between two recently separated eco-morphologically divergent populations (Malinsky et al. 2015). We explored divergence in the cranial lateral line by quantifying pore size in four of the canals using evidence from micro computed tomography scans. We also quantified divergence in the number of trunk canal and superficial neuromasts found across the head, body and fins using images of stained neuromasts. After accounting for variation linked to body size and sex, we found the deep water benthic ecomorph had larger cranial canal pores than the shallow water littoral ecomorph (Figure 2.4). However, we found no difference between the ecomorphs in number of trunk canal or superficial neuromasts on the head, body and fins (Figure 2.5).

2.6.1 Divergence of cranial canal morphology

Our observation of larger canal pores in the benthic ecomorph than littoral morph is suggestive of divergence in their sensory systems more broadly, and hence functional differences in their ability to detect hydrodynamic stimuli. The positioning of neuromasts in canals is believed to

enable fishes to reliably detect alternating current stimuli in the environment (Bleckmann and Zelick, 2009; Klein and Bleckmann, 2015), for example of the form generated by fish or zooplankton prey in the environment, or generated by hidden motile prey hidden in the muddy or sandy substrates (Schwalbe et al. 2012). The presence of larger canal pores is likely to be associated with a greater reliance on the lateral line system for survival, for example in more turbid or darker conditions (Marshall, 1996; Maranzino and Webb, 2018). There have been no measurements of turbidity in the Lake Masoko, which may vary by depth, but the intensity of light (Figure 2.1C) and range of wavelengths of light both decline substantially by depth, and deeper water fish have experienced selection on their visual system in line with expectations from the measured light environment (Malinsky et al. 2015; Carruthers et al. in prep). It is possible, therefore, that in the benthic zone, visual cues have a reduced role in communication with conspecifics, in behaviours such as schooling for predator avoidance, male-male competition for breeding territories, and male-female interactions during mating, and that this is compensated for by an enhanced role for lateral line-mediated detection of key signals (e.g. Faucher et al. 2010; Butler and Maruska 2015).

The observed divergence in cranial canal morphology may be additionally linked to the differing diets of the two ecomorphs. Stable isotope analysis has revealed that the muscle of the littoral ecomorph is relatively enriched in ^{13}C, reflective of a diet dominated by littoral arthropods, while muscle of the benthic ecomorph is relatively depleted in ^{13}C, indicating a more offshore zooplantivorous diet (Malinsky et al. 2015; Carruthers et al. 2022) (Figure 2.1E). Previous research has identified that dietary grouping can be a useful predictor of cranial lateral line morphology across the Lake Malawi radiation (Edgley and Genner, 2019). However, we found no consistent association between pore size and trophic niche when comparing zooplankton feeders (i.e. the "utaka" and *Diplotaxodon* groups of species) to species feeding on bottom-living invertebrates (i.e. the "shallow benthic" and "deep benthic" species groups) (Figure 2.6). Instead, inside the main Lake Malawi radiation, it is predominantly the case that the species with diets dominated by substrate-living invertebrate prey tend to have the largest pores, particularly molluscivorous species (Edgley and Genner, 2019) (Figure 2.6). Notably, however, there is a consistent pattern that where species groups with similar diets living at different depths are compared, those living in the deeper waters have larger cranial canal pores (e.g. shallow living "utaka" vs. deep living *Diplotaxodon* groups, and the "shallow benthic" vs. "deep benthic" groups) (Figure 2.6A-D). This may suggest that constraints imposed by both the dark light regime and the requirement to detect motile prey may combine to drive divergence in observed lateral line system phenotype.

2.6.2 Canal and superficial neuromast divergence

Similar to cephalic canals, trunk canals are engaged with sensing alternating current flow, so may be best able to detect flow generated by proximate fish or prey items (Klein and Bleckmann, 2015). In contrast, superficial neuromasts are thought to be primarily influenced by direct current, including any flow in the environment, enabling behaviours such as rheotaxis, but they are also useful for sensing movement in non-flow environments, particularly where visual cues may be limited. Many deep-sea fishes, for example, have an expanded number of superficial neuromasts on the head and body (Marranzino and Webb, 2018). Similarly, the blind ecomorph of the Mexican tetra *Astyanax mexicanus* that inhabits cave environments where light and current are both absent, have vastly increased numbers of superficial neuromasts compared to congenerics in surface habitats (Yoshizawa et al. 2014). Thus, we may expect the benthic ecomorph of Lake Masoko *A. calliptera* to have more trunk canal neuromasts and superficial neuromasts on their body due to deeper and darker environment. However, we found no evidence of divergence between the ecomorphs in the number of trunk canal or superficial neuromasts (Figure 2.5B-C). This may suggest that divergent light environments are not driving sufficiently strong selection to lead to divergent adaptation of the trunk canal or superficial components of the lateral line within Lake Masoko.

In contrast to the lack of divergence observed in neuromast counts of Lake Masoko fish, we found substantial variation in neuromasts numbers of the wider Lake Malawi radiation. Here we found that the number of superficial neuromasts was broadly associated with life history (Figure 2.6). The typically open water feeding *Rhamphochromis*, *Diplotaxodon* and "utaka" clades, and the demersal feeding "shallow benthic" and "deep benthic" clades had greatest number of neuromasts on average in the trunk canals (Figure 2.6E-H). Conversely, "mbuna" species inhabiting the rocky-shore littoral zone of Lake Malawi, that predominantly feed on epilithic algae and similar resources (Ribbink et al. 1983; Genner and Turner 2005) generally have few trunk canal or superficial neuromasts (Figure 2.6E-H). Notably, *A. calliptera* had a similar number of neuromasts to the mbuna species. Although typically omnivorous, it is also phylogenetically nested within the mbuna clade (Malinsky et al. 2018), perhaps indicative of phylogenetic constraint on lateral line phenotypes.

2.6.3 Selection and constraints on lateral line divergence

The evidence of divergence in lateral line phenotypes in Lake Masoko over a timescale of less than 1,000 years (Malinsky et al. 2015) (Figure 2.4), combined with evidence of divergence across the

broader Lake Malawi radiation that has evolved over the last one million years (Malinsky et al. 2018) (Figure 2.6), is consistent with a role for natural selection in shaping lateral line phenotypes, and promoting broad-scale evolutionary divergence of these cichlid fishes. However, our results raise multiple issues that will require further work to address. Notably, our analyses have not confirmed a genetic basis to the observed cranial canal line variation in Lake Masoko *A. calliptera*. It is possible observed phenotypic variation has arisen from developmental plastic responses to differing resource availability. Confirmation of a fixed genetic basis to the divergence will require quantification of the morphology of fish from each genetic background that have been reared in common garden conditions. Additionally, it may be possible to identify genetic variants associated with the trait in either hybrids (e.g. Navon et al. 2021) and/or by studying expression quantitative trait loci (eQTLs) (e.g. Carruthers et al. 2022). Linking evidence from single nucleotide polymorphisms (SNPs) with expression data from wild Lake Masoko *A. calliptera* has proved successful for identification of key functional genes under selection that influence lower pharyngeal jaw shape, including documenting a role for genes associate with bone development (Carruthers et al. 2022; Vernaz et al. 2022)

Our results do not enable us to conclusively identify the key drivers of divergence in the lateral line phenotypes of Lake Malawi haplochromines. Like visual systems, it is likely that multiple ecological factors are mediating selection, including diet, habitat and social behaviours that can vary across ontogeny (Montgomery et al. 2014). There are also likely to be phylogenetic constraints that limit standing genetic variation and therefore the phenotypes that selection can generate (McKitrick 1993; Edgley and Genner, 2019). Importantly, cephalic lateral line phenotypes are related to other key aspects of head morphology, including jaw, opercular and eye morphology (Parsons and Albertson, 2006; Bird and Webb, 2014; Edgley and Genner, 2019). Thus, there appear to be intrinsic constraints on the types of lateral line canal structures that can develop in the context of key aspects of morphology. It is possible, for example, that the relatively larger head of the deep-water benthic Lake Masoko ecomorph (Figure 2.1G) can reliably accommodate the larger pores, that would not be feasible in the shorter jawed shallow water littoral ecomorph. Closer explorations of the covariance between lateral line canal structures and broader aspects of craniofacial morphology, both within and between ecomorphs, coupled with studies of development (Baker et al. 2013; Bird and Webb, 2014; Becker et al. 2016), would help to resolve constraints and modularity of the system.

2.6.4 Conclusions

There is now a wealth of evidence that ecological speciation in East African cichlids is enabled by adaptation to different habitats and trophic niches (Danley and Kocher, 2001; Ronco et al., 2021), which in turn are facilitated by the evolution of divergent sensory systems – e.g. vision (Carleton et al. 2008) – and ecomorphological traits – e.g. cichlid pharyngeal and oral jaws (Galis and Drucker, 1996; Seehausen 2006. In the Lake Masoko system we are fortunate in being able to observe the early stages of a speciation event, including clear evidence of niche partitioning of populations that is accompanied by lateral line divergence, in addition to more commonly studied visual systems and ecomorphological phenotypes. Our results suggest the Lake Masoko system may provide opportunities to explore the evolution of the cranial canal system in natural ecological context. There are clear opportunities for the system to be studied experimentally to learn more about how regulation of gene expression changes during development, and how lateral line morphology interacts with behaviour, for example schooling (Partridge and Pitcher, 1980; Faucher et al. 2010) and male-male competition (Butler and Maruska, 2015). We suggest that the lateral line system is a vital component of the functional morphology of fishes, which requires detailed consideration if we are to better understand mechanisms driving adaptive radiation in fishes.

2.7 Supplementary Information

Table S2.1. Models of both neuromast counts and pore size measurements, comparing Lake Masoko with the Lake Malawi radiation (Figure 2.6). For the four pore area analyses, 'group' indicates the Lake Malawi clade (mbuna, utaka, *Diplotaxodon*, *Rhamphochromis*, shallow benthic or deep benthic), and Masoko subpopulation (littoral, intermediate and benthic). For the four neuromast count analyses, 'group' indicates Lake Malawi clade (mbuna, utaka, *Diplotaxodon*, *Rhamphochromis*, shallow benthic, deep benthic or *A. calliptera* 'Salima'), and Lake Masoko capture depths (deep or shallow).

Family	Response	Predictor	Estimate	p-value
gaussian	Log_{10} mean mandibular pore area	Group	82.74	**< 0.001*** **
		Log_{10} standard length	166.54	**< 0.001*** **
gaussian	Log_{10} mean preopercular pore area	Group	73.59	**< 0.001*** **
		Log_{10} standard length	109.43	**< 0.001*** **
gaussian	Log_{10} mean infraorbital pore area	Group	38.40	**< 0.001*** **
		Log_{10} standard length	89.08	**< 0.001*** **
gaussian	Log_{10} mean otic pore area	Group	30.58	**< 0.001*** **
		Log_{10} standard length	132.98	**< 0.001*** **
poisson	Log_{10} total number of head superficial neuromasts	Group	96.16	**< 0.001*** **
		Log_{10} standard length	15.59	**< 0.001*** **
poisson	Log_{10} total number of head superficial neuromasts	Group	34.979	**< 0.001*** **
		Log_{10} standard length	0.552	0.459
poisson	Log_{10} total number of head superficial neuromasts	Group	5.383	**< 0.001*** **
		Log_{10} standard length	0.638	0.427
poisson	Log_{10} average superifical neuromasts per trunk cluster	Group	21.395	**< 0.001*** **
		Log_{10} standard length	0.057	0.812

An *Aulonocara stuartgranti* × *Astatotilapia calliptera* F_2 hybrid
Photo: Emília Santos

Chapter 3

Genome-wide association studies reveal loci associated with lateral line system diversity in haplochromine cichlids

3.1 Abstract

The lateral line system is a mechanoreceptive sensory system in fishes which is vital for many behaviours including detecting prey, orientating to water flow and avoiding predators. In the Lake Malawi haplochromine cichlid radiation the system has diverged extensively in morphology, linked to diet and environmental niche. Despite its ubiquity in fishes and its adaptive importance, the genomic basis for lateral line system disparity has not been extensively studied. Here, we produced an F_2 cross of two species of Lake Malawi cichlid with divergent lateral line system phenotypes, finding clear segregation of morphology among hybrids. We found that for both cephalic canal pore morphology and neuromast number, hybrids exhibited extensive variation, which in most cases was intermediate between the phenotypes of the two parental populations. Genome-wide association studies (GWASs), testing for associations between genomic variants and our lateral line system phenotypes, revealed candidate genes influencing lateral line system disparity, including genes known to be associated with craniofacial development (*bmp4*) and neuromast function (*nsfa*). These results give the first insights into the genomic basis of lateral line system diversification during cichlid adaptive radiation, and provide a basis on which to further investigate the genetic origins of diversity in this often overlooked sensory modality.

3.2 Author Contributions

Duncan Edgley conducted all lateral line system imaging, phenotyping and analysis, conducted all GWAS and GO enrichment analysis, and wrote the manuscript. This chapter was conceived, conducted and written under the supervision of **Martin Genner**.

Samples were collected and the cross was bred by **George Turner**. Fish were fed and maintained by **George Turner**, **Michael Hayle**, **Emília Santos** and **Duncan Edgley**. Phenotyping was conducted in collaboration with another project phenotyping pigmentation, conducted by **Emília Santos**, with **Aleksandra Marconi**, **Bethan Clark**, **Aaron Hickey**, **Gregoire Vernaz**, **Tyler Linderoth** and **Melany Henot**. **Thomas Davies** and **Elizabeth Martin-Silverstone** assisted with CT scanning and related software. Whole genome sequencing was conducted at the Sanger Institute. Read mapping and variant calling was conducted with contributions from **Richard Durbin**, **Bettina Fischer**, **Jake Morris**, **Aaron Hickey** and **Moritz Blumer**. GWAS and GO enrichment was conducted with assistance from, and using adapted code from **Tyler Linderoth** and **Madeleine Carruthers**. A particular contribution to this chapter was made by both **George Turner** and **Emília Santos** throughout the 9 phenotyping trips to Bangor over the past 4 years.

3.3 Introduction

In adaptive radiations morphologies tend to be closely related to function, having adapted to the specific requirements of their niche (Albertson et al. 2003; Edgley and Genner 2019; Ronco et al. 2021). As such, research has often focused on understanding the functional consequences of variation in phenotypes and correlating these with species ecology. However, a key aspect of understanding the dynamics of how species diversify, is awareness of the genetic basis of observed functional morphological diversity (Irschick 2002; Albertson et al. 2003; Rundle and Nosil 2005), which can be obtained by mapping phenotype to genotype (Albertson et al. 2018). Current approaches for this, which have been used successfully for identifying of candidate loci associated with functional phenotypes in fish, include genome-wide association studies (GWAS) (Munby et al. 2021), quantitative trait locus (QTL) mapping of hybrid crosses (Albertson et al. 2003; Parsons et al. 2012; Wark et al. 2012), and RNAseq-based gene expression analyses (Ahi et al. 2019; Carruthers et al. 2022).

Sensory systems inform multiple behaviours vital for survival (Sabbah et al. 2010; Kermen et al. 2013; Mogdans 2019), and diversification of these sensory systems may contribute to speciation. For example, fish visual systems tend to be adapted to the light regime of their environments (Carleton et al. 2005; Sabbah et al. 2010) and are fine-tuned during evolution, and hence may promote reproductive isolation (Seehausen et al. 2008). However, despite this, not all sensory modalities have been thoroughly investigated. The lateral line system of fishes, for example, has been found to be fine-tuned to specific trophic (Edgley and Genner 2019) and/or habitat niches (Wark and Peichel 2010; Marranzino and Webb 2018), but knowledge of genomic basis of lateral line diversity is limited to only some aspects of lateral line diversity, and only in three-spined sticklebacks (Wark et al. 2012).

All living fish species possess a functional mechanoreceptive lateral line system, allowing them to detect flow of water around them (Bleckmann and Zelick 2009; Mogdans 2019). Fishes use their lateral line systems for a variety of functions, including navigation (Montgomery et al. 1997; Suli et al. 2012), shoaling (Faucher et al. 2010), finding prey (Montgomery and MacDonald 1987; Schwalbe et al. 2012), avoiding predators (Stewart et al. 2014), and intraspecific communication (Butler and Maruska 2015; Mogdans 2019). Consequently, selection is likely to have modified the lateral line morphologies of species to the requirements of their respective environments, implying that it is a major axis of adaptation during fish population diversification. This is particularly evident in the cichlid adaptive radiations of the East African great lakes, in which a myriad of

lacustrine species have evolved from riverine ancestors over relatively short periods of evolutionary time (Malinsky et al. 2018; Kautt et al. 2020; Ronco et al. 2021). In the Lake Malawi radiation, for example, cichlid species have evolved a remarkable degree of variation in lateral line phenotypes, both in neuromast patterning and skeletal morphology, corresponding to trophic and habitat niches (Webb et al. 2014; Edgley and Genner 2019; see also Chapter 2), within the past million years (Genner et al. 2007; Malinsky et al. 2018).

Phenotype-genotype mapping through GWAS or QTL mapping analyses has been successful in identifying candidate genes associated with diversification in craniofacial shape in cichlids. The shape of the oral and pharyngeal jaws, and the shape of the head more generally has diversified dramatically during cichlid adaptive radiation (Albertson et al. 2003), as an adaptation to the different trophic niches in the Great Lakes (Muschick et al. 2012; Edgley and Genner 2019; Conith and Albertson 2021). In particular, QTL mapping has found loci known to be involved in determining head shape across many taxa (Streelman and Albertson 2006), and gene expression analyses have identified regions under selection during population divergence (Ahi et al. 2019; Carruthers et al. 2022).

Despite these advances, with the exception of vision (Carleton et al. 2005; Hofmann et al. 2009; Fan et al. 2012; Malinsky et al. 2015), research on the genomic changes that underpin variation in adaptive sensory system morphology is lacking. For example, these methods have not yet been applied to the full diversity of lateral line system morphology in cichlid fishes. Initial research has been carried out in three-spined sticklebacks, where QTL mapping methods have identified candidate loci associated with neuromast patterning, in particular highlighting the significance of the *Eda* gene (Wark et al. 2012; Ahnelt et al. 2021). Specifically it was found that multiple linked traits, including lateral bony plate architecture, superficial neuromast number, and even schooling behaviour can be mapped to variation near the *Eda* locus (Greenwood et al. 2016; Archambeault et al. 2020; Wagner et al. 2022). These initial findings have the potential to form the basis for investigating the diversity of lateral line system morphology in a range of taxa, including rapidly diversifying clades such as cichlids.

In determining the genomic basis for diversity in such a complex and integrated system as lateral line system morphology, many difficulties remain. Firstly, the lateral line of fishes is modular, and has multiple integrated components (i.e. the canal and superficial lateral lines), each of which has different neuromast structures (Windsor and McHenry 2009; Webb 2014), and distinct functions

(Bleckmann and Zelick 2009; Mogdans and Bleckmann 2012). In addition, similar to craniofacial architecture in cichlids, the genetic basis of lateral line diversification is likely to be complex and highly polygenic (Albertson et al. 2003; Mills et al. 2014), requiring a high resolution to detect associations when detecting phenotype-genotype interactions.

Here, we present a study mapping phenotype to genotype across several lateral line system phenotypes in Lake Malawi cichlids. We focus on two species of cichlid from the Malawi radiation which have diversified significantly in their lateral line system morphologies (Edgley and Genner 2019). Through breeding an F_2 interspecific cross, and correlating phenotypic and genotypic variation among hybrids, we can gain insights into both the degree of heritability of these traits, but also identify potential genomic loci involved in adaptation of the lateral line system during adaptive radiation in Lake Malawi. One parent species of this cross, *Aulonocara stuartgranti*, commonly known as the peacock cichlid, is a sediment-feeder, most commonly seen inhabiting the rock-sand interface habitats of Lake Malawi at depth between 5 and 15m (Konings 2016). It has been observed in the wild "probing" the sandy or muddy substrate with its lateral line system to hunt for cryptic prey (Konings, 2016). In laboratory trials it has been shown to be able to use its canal lateral line system to detect prey movement in the complete absence of light (Schwalbe et al. 2012; Schwalbe and Webb 2015). It has evolved dramatically widened canals and pores of the cephalic lateral line system (Webb et al. 2014; Becker et al. 2016; Edgley and Genner 2019) (Figure 3.1A, 3.1C, 3.1E; Figure S3.1A), which has led the species to be referred to as a "sonar feeder" (Konings, 2016). For the other parent of the cross, we chose *Astatotilapia calliptera*, a cichlid found in the shallow water of Lake Malawi and its surrounding rivers and satellite lakes (Parsons, Márquez, and Albertson 2012; Konings, Ad 2016; Malinsky et al. 2018). Living in these more heterogeneous environments, it has experienced more environmental variation between wet and dry seasons (Delalande, 2008; Malinsky et al. 2015), and less consistent sources of food (Turner et al. 2019) during its evolution. It has a less specialised lateral line morphology, with narrow canals, and relatively few superficial neuromasts (Edgley and Genner 2019) (Figure 3.1B, 3.1D, 3.1F; Figure S3.1B).

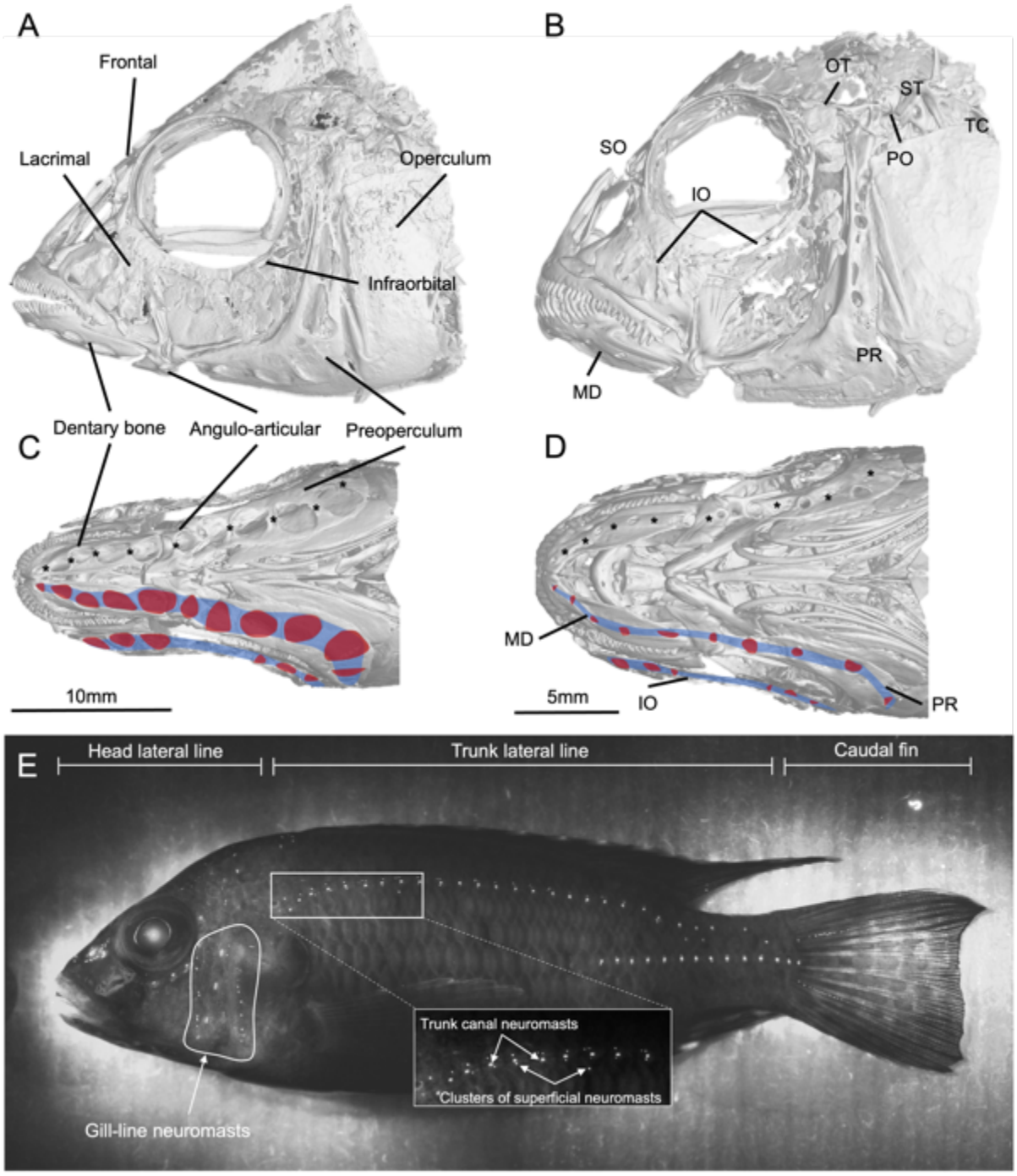


Figure 3.1 The lateral line system of the two parental species included in this study: *Aulonocara stuartgranti* and *Astatotilapia calliptera*. (A) Lateral view of the widened cephalic morphology of *A. stuartgranti*. (B) Lateral view of the narrow cephalic morphology of *A. calliptera*. (C) Ventral view of the cephalic morphology of *A. stuartgranti*. (D) Ventral view of the cephalic morphology of *A. calliptera*. (E) Definitions of the neuromast counts included in this study, illustrated on an F_2 hybrid. Asterisks indicate approximate locations of canal neuromasts. CT scans are annotated with the names of the cranial bones which contain lateral line canals. MD = mandibular canal; IO = infraorbital canal; PR = preopercular canal; SO = supraorbital canal; OT = otic canal; PO = postotic canal; ST = supratemporal canal; TC = trunk canal.

By crossing these two species with highly divergent lateral line system morphologies, and conducting genome-wide association studies (GWAS), we can begin to understand the genetic basis for lateral line system disparity, in cichlids and across other clades of fishes. We aim for this research to provide a basis on which we can further investigate the genetic changes that underpin the remarkable diversity in lateral line system morphology we observe.

3.4 Materials and Methods

3.4.1 Cross design and Sample collection

We used individuals of two species of Lake Malawi cichlid chosen for their highly divergent lateral line system morphologies: the lacustrine "sonar-feeder" *Aulonocara stuartgranti* and the generalist *Astatotilapia calliptera* (Figure 3.1). The male F_0 founder of the cross was a male *A. stuartgranti*, sourced from an aquarium strain established in 2011. The strain was originally wild-caught from from Usisya, a small village north of Nkhata Bay in Malawi (11°14'16.8"S 34°13'58.8"E) and was kept inbred prior to establishment of the cross. *A. calliptera* females were wild-caught and sourced as fry from Itupi stream in south-western Tanzania (9°20'15.0"S 33°45'03.6"E) in 2015 and 2016. We selected several females from this strain to cross with the *A. stuartgranti* male founder.

We produced two broods of F_1 hybrids from the two *A. calliptera* mothers, which were subsequently pooled into one population. Random inbreeding between these F_1 hybrids produced multiple broods of F_2 hybrids (n=124) which were similarly isolated and pooled in tanks. The process was repeated for F_3 individuals (n=24). All crosses were bred in a consistent environment of 120 litre tanks and fed on a mixture of flake and pellet. We only included males in our analysis. We ensured that all hybrid individuals chosen for phenotyping were fully mature and exhibiting male breeding colouration. In addition, we ensured that the standard length of each individual was at least 5cm.

3.4.2 Lateral line system phenotypes

To quantify lateral line system disparity among our F_2 and F_3 individuals and their parents, we used microCT scanning to visualise the cranial canals of the head. We CT scanned all fish in a Nikon XTH225ST micro-computed tomography (micro-CT) system at the University of Bristol. Scans of fish heads were conducted in a similar manner to as outlined in Chapter 2, and resolution was also comparable (voxel size 20-25 μm). Each individual scans involved 3141 projections, and included two preserved fish. Following scanning and reconstruction, we produced a 3D model of each individual in VGStudio MAX 3.3.6 (Volume Graphics GmbH 2016) and captured 2D still

images of the lateral view and the ventral view of the jaw for all individuals. Similarly to previous work, we phenotyped the mandibular canal, the infra-orbital canal from the ventral perspective, and the preopercular and otic canals from the lateral perspective. Using tpsDig and tpsUtil (Rohlf 2015), we estimated the area of each pore of the aforementioned canals. We drew curves of sliding semilandmarks around the circumference of each pore on our scans, bounded by a landmark at either end (at the same location) at the anterior limit of the pore. These curves were resampled to nine equidistant points, and converted to landmarks (Rohlf 2015). Pore area was estimated by considering the first nine points of each pore's curve as the vertices of a polygon, and calculating the area of this polygon in the geometry package in R 4.2.0 (Adams et al. 2022; Roussel et al. 2022). In total, we microCT scanned 124 F_2 individuals, 24 F_3 individuals, three *Aulonocara stuartgranti* from the "Usisya" lab strain, and three *Astatotilapia calliptera* wild caught from Itupi stream.

To visualise the neuromasts of each individual, we placed each live fish in 0.008% DASPEI solution [2-(4-(dimethylamino)styryl)-N-ethylpridinium iodide; Fisher Scientific]. DASPEI is a vital mitochondrial dye commonly used to stain the mitochondria of live hair cells, such as those of the neuromast. It has been used successfully previously in closely related species (Becker et al. 2016; Butler et al. 2016). Following 30 minutes of submersion in the dye, fish were euthanised in MS-222 and death confirmed by exsanguination. Photographs of the left side of the fish were taken using a Canon EOS-500 DSLR with a Sigma 18-200 mm f/3.5-6.3 lens, under a yellow longpass filter (500 nm) and a royal blue UV lamp. For each fish we took multiple individual images, highlighting different parts of each fish's lateral line system. Images were stitched together in ImageJ and adjusted to increase contrast and aid visualisation of neuromasts (Schneider et al. 2012) (Figure 3.1E). Neuromasts were counted only on the left side of the fish, though for the underside of the jaw, we counted neuromasts on both sides. We found no evidence for any significant differences in the numbers of neuromasts on the left and right sides of the body, validating our choice to only count neuromasts on the one side of each specimen both here and in other morphometrics studies included in this thesis (Figure S3.2). In total, we DASPEI stained and imaged 124 F_2 individuals, 24 F_3 individuals, three *Aulonocara stuartgranti* from the "Usisya" lab strain, and seven *Astatotilapia calliptera* from Itupi stream.

3.4.3 Genotyping

DNA was extracted from pectoral fin-clips for each individual and fish were whole-genome shotgun sequenced using an Illumina HiSeqX. F_0 parental founders were sequenced at ~35×, while F_2 and F_3 individuals were sequenced at an average of ~12.5× coverage. We mapped reads to the

fAstCal1.2 *Astatotilapia calliptera* reference genome (accession GCF_900246225.1) with bwa-mem 0.7.17 (Li 2013). Variants were called and normalised with bcftools mpileup and norm respectively (Li 2011). We then merged all multiallelic sites, merged chromosomes and removed non-biallelic SNPs. We conducted quality filtering using a combination of custom scripts and bcftools (Li 2011), to remove SNPs with > 25% missing data and a mean GQ of <30. Additionally, and prior to any genome-wide analyses, we followed steps in Munby et al. (2021) for quality filtering in gatk 4.2.3 (Poplin et al. 2018) and vcfCleaner (https://github.com/tplinderoth/ngsQC/tree/master/vcfCleaner). In contrast to the analyses by Munby et al. (2021), however, we excluded SNPs with a minor allele frequency of <0.1, and removed SNPs in linkage disequilibrium (LD) using plink 2.00 (Chang et al. 2015). Our LD filters were relatively loose, given the extent of shared genomic variation between our closely related individuals. Using strict LD filtering for our dataset resulted in large numbers of potentially informative SNPs being removed. After filtering, 3,752,375 SNPs were used for genomic principal component analysis (PCA) in plink 2.00 (Chang et al. 2015) (Figure S3.3).

3.4.4 Genome wide association studies

We chose 11 lateral line system phenotypes to test for associations with the genome through GWAS: mean mandibular canal pore area; mean preopercular canal pore area; mean infraorbital canal pore area; mean otic canal pore area (Figure 3.1A-D); total number of head superficial neuromasts; total number of gill-line superficial neuromasts; total number of trunk canal neuromasts; total number of trunk superficial neuromasts; average number of trunk superficial neuromasts per canal neuromast cluster; total number of lower jaw superficial neuromasts; and the total number of caudal superficial neuromast lines (Figure 3.1E). For each variable we excluded any individuals with missing data, or outlier individuals with clear extreme values (e.g. Figure 3.3C, D & G).

For each GWAS, we filtered our quality-filtered callset to only F_2 and F_3 individuals in bcftools (Quinlan and Hall, 2010). Each GWAS was conducted in GEMMA 0.98.1 (Zhou and Stephens 2012). This involved regression of each lateral line phenotype against posterior mean genotype as calculated from genotype likelihoods from our variants, using linear mixed models. We included a pairwise kinship matrix in our analysis as a random effect to account for the effect of shared ancestry, calculated in GEMMA 0.98.1 (Zhou and Stephens 2012). To account for the confounding effects of population stratification (Figure S3.3), PC1 – PC5 of the genomic PCA were included as covariates in all linear mixed models (LMMs) (Zhao et al. 2018). We also included

standard length of each individual as a covariate in each GWAS to account for the known effects of ontogeny on lateral line system morphology (Webb et al. 2014; Montgomery et al. 1994).

SNPs significantly associated with traits were identified using p-values from a likelihood ratio test on each LMM, included as output in GEMMA (Zhou and Stephens, 2012). To identify those significantly associated SNPs for subsequent gene ontology analysis, we used a p = 0.00001 threshold. Any p-values lower than this threshold were considered for future analyses. We pooled all significant SNPs from the four cranial canal GWASs (n=92) and the seven neuromast count GWASs (n=290) and used the "closest" feature in bedtools 2.28.0 (Quinlan and Hall, 2010) to identify the nearest gene to each of our associated SNPs, aligning against the fAstCal1.2 reference genome gene coordinates. We excluded any SNPs where the closest gene was > 10,000 base pairs away.

3.4.5 Gene ontology term enrichment analyses

We pooled our lists of candidate genes from GWAS and compared them against a list of 29,243 reference genes from the annotated *A. calliptera* genome. We annotated each gene with biological process gene ontology (GO) terms by comparing against a reference list of UniProt *Danio rerio* annotations aligned in bedtools 2.28.0 (Quinlan and Hall 2010). We obtained p-values for our GO terms using a Fisher's exact test in topGO 2.41.0 (Alexa and Rahnenfuhrer 2022), with an FDR correction of 0.01 for multiple testing. We used adapted functions from the package ViSEAGO 1.3.16 (Brionne et al. 2019) to cluster GO terms of genes based on Wang's measure of semantic similarity (Wang et al. 2007) and visualised these using multi-dimensional scaling plots (MDS). We computed these semantic similarity measures using the clustering algorithm in ViSEAGO (Brionne et al. 2019) for our set of significant GO terms calculated above. In addition, we conducted gene enrichment analyses for GO terms from the four cranial canal GWASs, and the seven neuromast count GWASs individually, as two additional tests. Here we used a looser FDR threshold of <0.05 to include a higher proportion of potentially informative SNPs from our GWASs.

3.5 Results

3.5.1 Segregation of lateral line morphology

For all four cranial canals examined in this study, we found clear segregation of morphology in the F_2/F_3 generations of *Aulonocara stuartgranti* × *Astatotilapia calliptera* hybrids (Figure 3.2). In all cases, *A stuartgranti* had the largest average pore area, *A. calliptera* had the smallest average pore area, and the hybrids were intermediate between the two (Figure 3.2). In terms of neuromast count, for the

most part we also observe the same trends. In some cases, such as the number of superficial neuromasts on the head, and the number of caudal lines of superficial neuromasts, the phenotype of the hybrid individuals exceeded that of both parents (Figure 3.3).

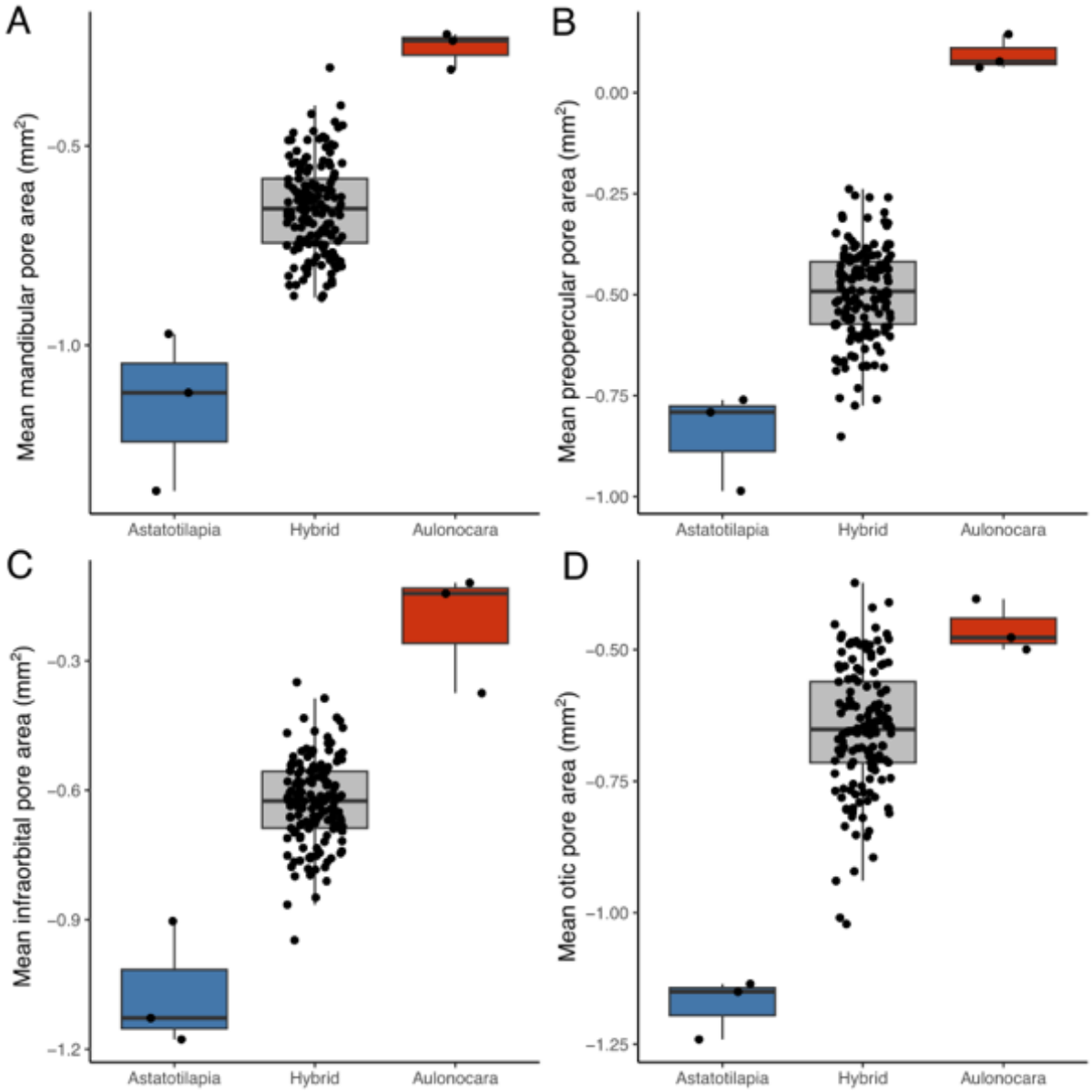


Figure 3.2 Pore areas for the four cranial canals of individuals included in this study. All plots are showing partialized residuals of the pore area measurement after accounting for standard length in a generalised linear model. (A) Mean mandibular canal pore area. (B) Mean preopercular canal pore area. (C) Mean infraorbital canal pore area. (D) Mean otic canal pore area.

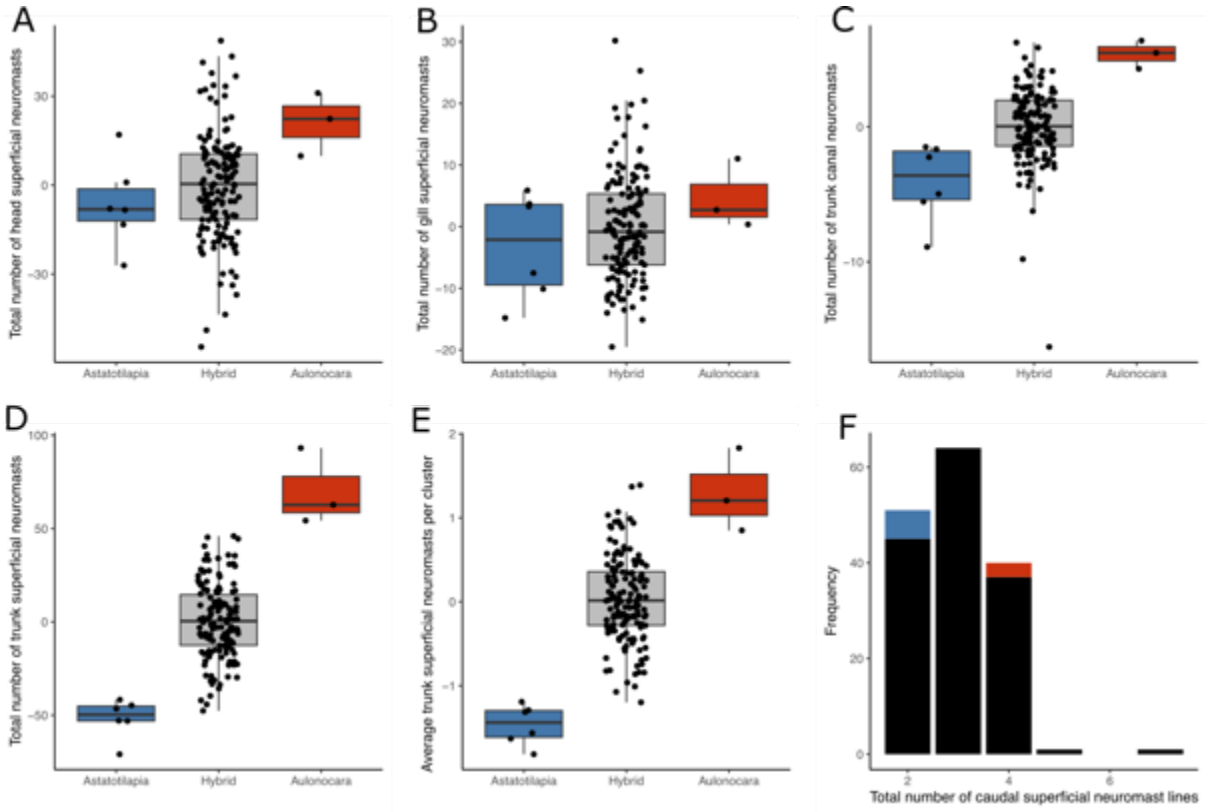


Figure 3.3 Variation in the numbers of superficial and canal neuromasts revealed by DASPEI imaging. A) The total number of head superficial neuromasts. (B) The total number of gill-line superficial neuromasts. (C) The total number of trunk canal neuromasts. (D) The total number of trunk superficial neuromasts. (E) The average number of trunk superficial neuromasts per trunk canal neuromast cluster. (F) The total number of lines of superficial neuromasts on the caudal fin.

3.5.2 GWAS and gene ontology term enrichment

We performed genome wide association studies (GWAS) for all eleven of our neuromast lateral line phenotype variables. Manhattan plots revealed distinct patterns of associations between genetic variants and phenotypes across the genome (Figure 3.4). Using SNPs exceeding the indicative 0.0001 p-value threshold from our GWAS, we looked for indications of the function of the closest genes to these SNPs using functional enrichment analysis of biological process gene ontology (GO) terms. When pooling the two GWAS groups and using a strict FDR threshold of 0.01, we found 46 GO terms enriched (Figure 3.5A-B). Clustering analysis, which groups SNPs by semantic similarity (Wang et al. 2007; Brionne, et al. 2019), revealed four clusters of biological process (bp) GO terms, relating to: neurogenesis and development; responses to oxygen-containing compounds; cellular processes; and metabolic processes (Table 3.1).

We also pooled genes into those associated with SNPs from the seven pooled neuromast count GWASs, and those from and the four pooled cranial canal GWASs. These resulted in 290 and 92 enriched GO terms respectively (Figure 3.6) using a relatively loose FDR correction threshold

(0.05). Including each list of genes separately in this clustering analysis (cranial canals and neuromast counts), revealed multiple GO terms for the cranial canals (Figure 3.6A), and neuromast counts (Figure 3.6B), covering genes ranging from cellular and metabolic processes to sensory systems. In both cases, there were relatively large clusters of genes involved in neural development. These clusters of genes have been annotated with GO terms of interest, which may be candidates involved in the disparity we observe in lateral line system morphology.

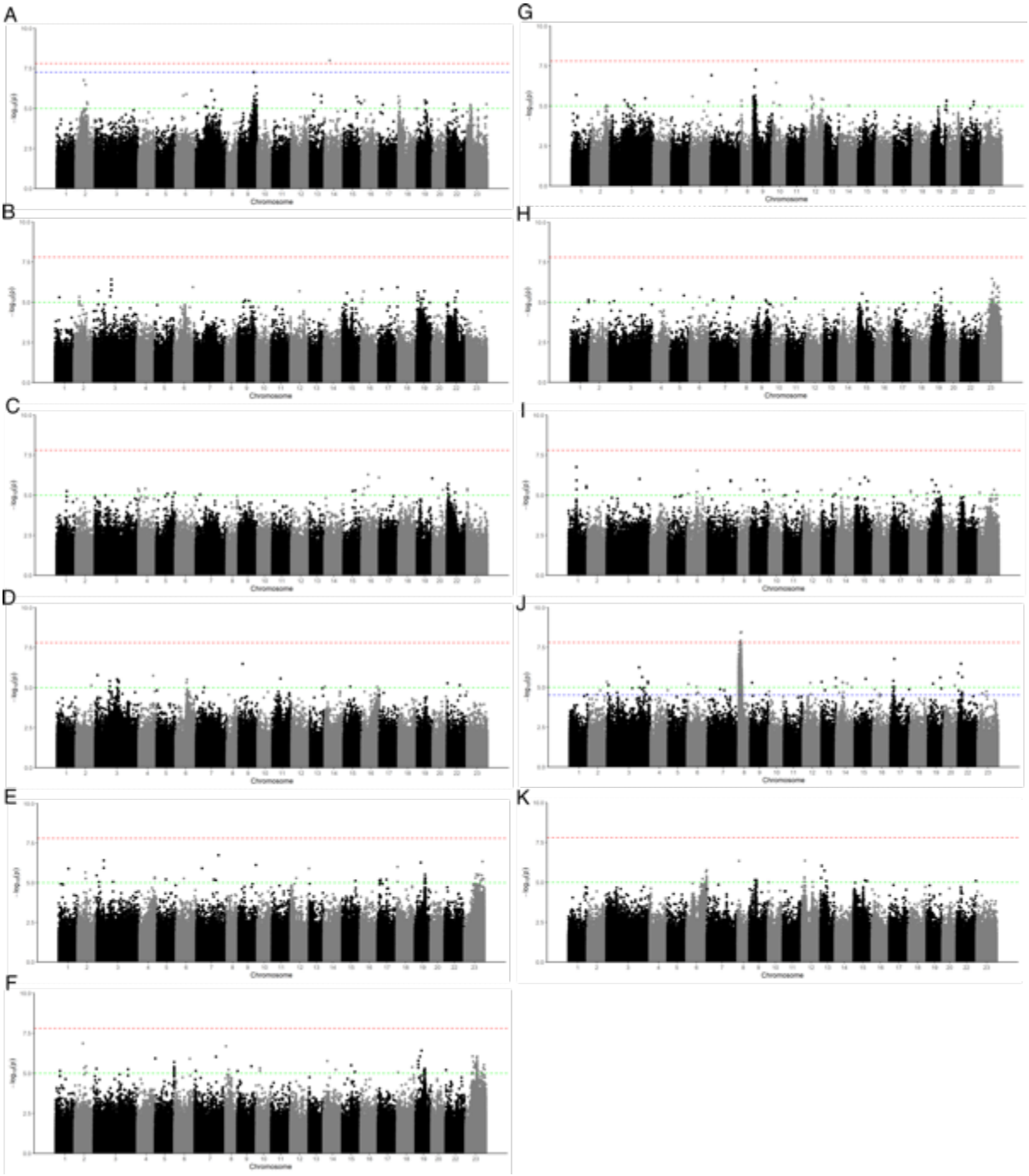


Figure 3.4 Manhattan plots from GWASs for (A-D) the average pore size of the four cranial canals of the lateral line system, and (E-K) neuromast counts. (A) Mean pore area for the mandibular canal. (B) Mean pore area for the preopercular canal. (C) Mean pore area for the infraorbital canal. (D) Mean pore area for the otic canal. (E) Total number of head superficial neuromasts. (F) Total number of gillline superficial neuromasts. (G) Total number of trunk canal neuromasts. (H) Total number of trunk superficial neruomasts. (I) Average number of trunk superficial neuromasts per canal neuromast cluster. (J) Total number of ventral superficial neruomasts. (K) Total number of caudal superficial neuromast lines. Green lines indicate the 0.00001 indicative threshold line, and red indicates Bonferroni correction threshold line. Blue indicates the 0.05 FDR correction threshold.

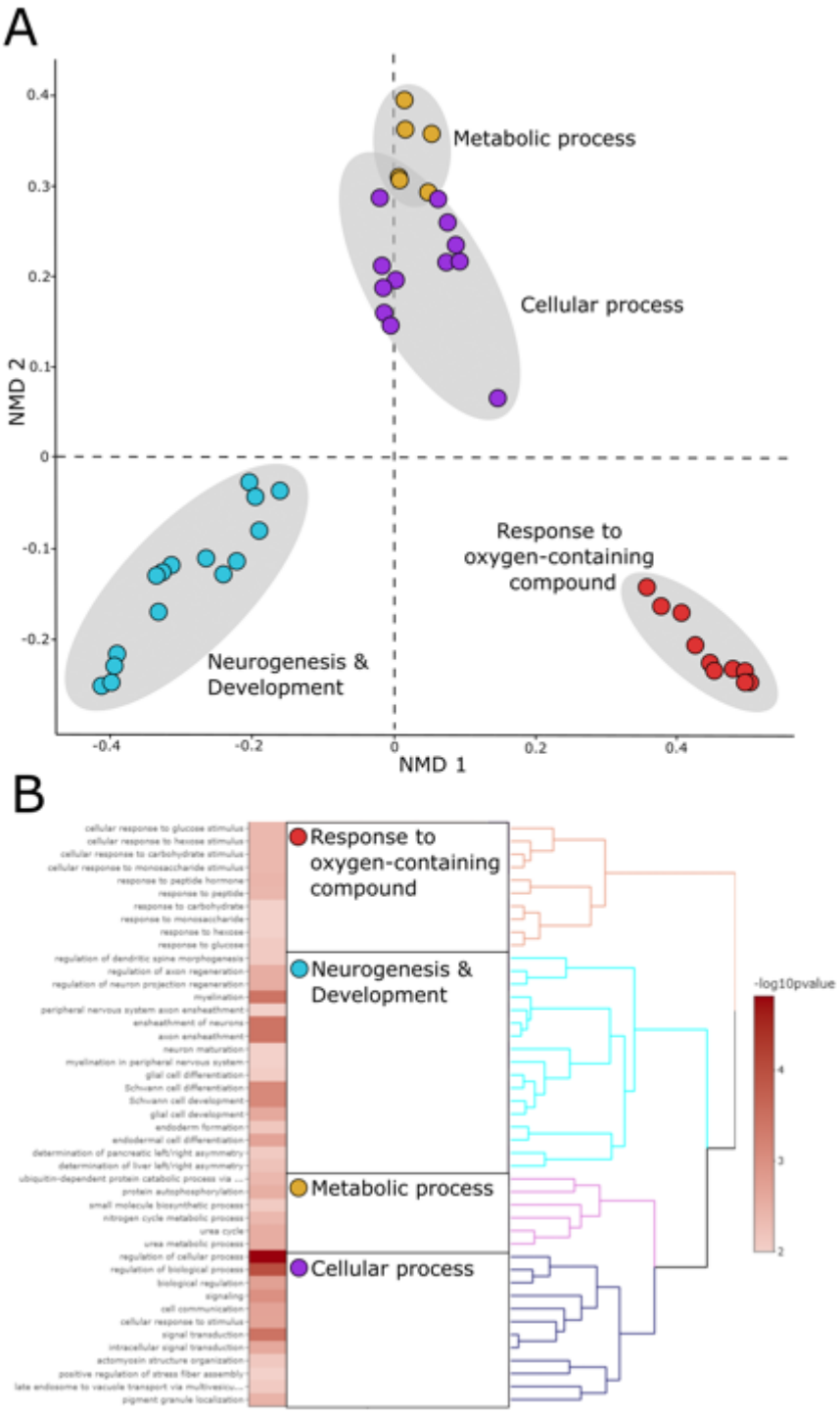


Figure 3.5 Significant GO terms associated with candidate genes, identified from a subset of SNPs from GWAS. Lists of genes were identified by finding the closest gene to each indicative SNP from GWAS by aligning to the annotated *A. calliptera* reference genome. (A) Multi-dimensional scaling (MDS) plot showing significant GO terms of biological processes (BP) following gene enrichment analysis in topGO, grouped into 4 GO clusters based on Wang's semantic similarity. Each circle is a single enriched GO term. (B) A list of significant GO terms from gene enrichment functional analysis in ViSEAGO, identified by Fisher's exact test with a FDR correction of <0.01. Right is a dendrogram, clustering GO terms based on Wang's semantic similarity distance into the four groups shown in A.

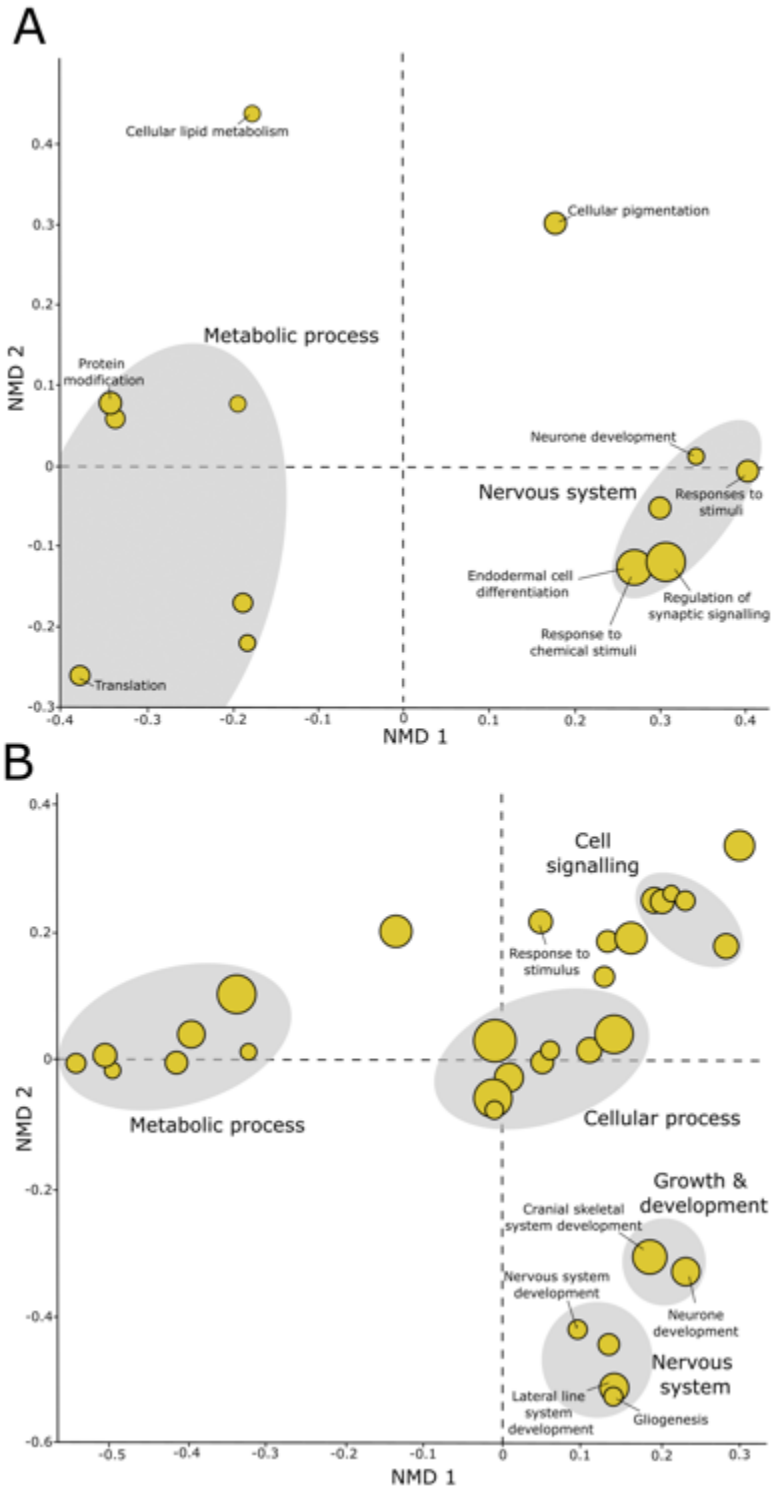


Figure 3.6 Clusters of significant GO terms, grouped by semantic similarity, for (A) the four pooled cranial canal GWASs, and (B) the seven pooled neuromast count GWASs. MDS plots show GO clusters calculated through Wang's semantic similarity in ViSEAGO, using gene enrichment tests with comparatively looser FDR correction than figure 6 (FDR<0.05). Each yellow circle is a GO cluster, with size proportional to the number of GO terms. Regions of similar GO clusters are bounded in grey, and the GO terms for candidate genes of interest for the lateral line system have been labelled.

Table 3.1 GO clusters and associated genes computed through gene enrichment analysis in topGO. GO clusters were computed based on semantic similarity of GO terms using Wang's method for subsets of candidate genes identified through GWAS. Shown are computed clusters for all 11 pooled GWASs at a FDR threshold of <0.01; GO clusters for the four cranial canal GWASs at FDR <0.05; and GO clusters for the seven neuromast count GWASs at FDR <0.05

Gene enrichment analysis	GO cluster	Number of GO terms in cluster	Selection of associated genes
All GWASs (FDR 0.01)	1 – Cellular process	10	pck2; stxbp4; grb7; pck2; stat1; crhr1
	2 – Neurogenesis and development	17	col12a1b; col12a1a; dync1h1; arhgef10; **nsfa**; fyna; il1rap1b; col12a1b; **bmp4**; myo1d
	3 – Metabolic process	6	vps37c; snf8; fyna; pim2; pimr23; pimr25; lipib; elovl6l; nags; pck2; glsa; asah2; sirt5; arg1
	4 – Response to oxygen-containing compound	12	pik3r6a; znf609b; reck; gapvd1; pdk2a; meis2a; ank3a; taar20v; gbp3; vta1; snf8; shroom2b; nsfa; rab27a
Cranial canals (FDR 0.05)	1 – Cellular lipid metabolic process	5	lipib
	2 – Metabolic process	5	clocka; larp1b; cpeb2; cnot7; gpc3; kdm4b
	3 – Nucleic acid metabolic process	6	cnot7
	4 – Protein modification process	7	kdm4b; ube2f; sirt5
	5 – Cellular amide metabolic process	5	larp1b; fars2; cpeb2; cnot7
	6 – Translation	6	fars2; larpb; cpeb2; cnot7
	7 – Cellular metabolic process	5	mgat5b; fars2; larp1b; cpeb2; cnot7; sirt5
	8 – Signalling	14	neto1l; grm8b; gpc3; gnal; taar20v; or131-2
	9 – Cellular pigmentation	7	rab27a
	10 – Transport	7	rab27a; gpc3; tnfaip2b
	11 – Neuron development	5	cc2d2a; col12a1b
	12 – Process	12	col12a1b; neb; fgfr3; or131-2; gnal; cc2d2a; si:zfos-588f8.1; cnot7
	13 – Process	7	clocka; col12a1b; cc2d2a
Neuromasts (FDR 0.05)	1 – Protein modification process	12	inhbb; **bmp4**; fyna; cdk5rap3; cdk5r1b; p4ha1a; pim2; pimr23; pimr25
	2 – Fatty acid biosynthetic process	5	elovl6l; asah2
	3 – Organic acid biosynthetic process	6	elovl6l; nags; glsa; asah2
	4 – Oxoacid metabolic process	8	glsa; gpt2l; elovl6l; nags; pck2; asah2; gpt2l; arg1
	5 – Catabolic process	8	arg1; nags; glsa; gpt2l; pck2; rnf213b; rragd
	6 – Metabolic process	15	pdk2a; ppp1r3cb; znk609b; pik3r6a; pck2; evol6l; asah2
	7 – Phosphorus metabolic process	5	pik3r6a; prkg1b; pdk2a; inhbb; dusp3a; map3k2; ducy2g; pld1a
	8 – Nitrogen compound metabolic process	10	vps37c; snf8; adam8b; plaub; elovl6l; arg1; nags; ctc1; asah2
	9 – Multicellular organ development	7	ahr2; asah2; pck2; myo1d; **bmp4**; pck2

10 – Developmental process	13	reck; il1rapl1b; **nsfa**; **bmp4**; col12a1a; lbh; meis2a; auts2a; cdk5rap3; fyna; lbh; myo1d
11 – Neuron development	10	il1rapl1b; **nsfa**; auts2a; col12a1a;
12 – Lateral line system development	10	**nsfa**
13 – Nervous system development	6	arhgef10; **nsfa**; dync1h1; fyna
14 – Gliogenesis	7	dync1h1; arhgef10; **nsfa**;
15 – Response to stimulus	8	pik3r6a; rnf214b; reck; gapvd1; syde1; gucy2g; gpr4; ahr2; ahr1b; sumo1; grb7; rragd; pck2
16 – Response to stimulus	8	rnf213b; fyna; stat1a; gpr4; adam8b; sost; mbp4; grb7; pck2; crhr1; pik3r6a; reck
17 – Response to oxygen-containing compound	9	pck2; stxbp4; rragd; crhr1
18 – Response to hormone	6	grb7; pck2; stat1a; stxbp4; crhr1; grb7
19 – Response to organonitrogen compound	11	grb7; pck2; stat1a; stxbp4; crhr1; grb7; rragd
20 – Protein localisation	14	**nsfa**; shroom2b; tango2; stxbp4; vps37; snf8; ank3a; rab40b
21 – Membrane repolarisation	6	kcnh2a
22 – Transport	11	vta1; snf8; vps37c; slc16a12b; dpp6a; kcnh2a; kcnj2a; kcnj16; kcnq5b; il1rapl1b; stxbp4
23 – Signal transduction	9	rnf213b; reck; sost; cdk5rap3; dusp3a; styxl2; **bmp4**; pik3r6a; inhbb
24 – Regulation of signal transduction	6	pik3r6a; rnf213b; reck; arhgef10; gpr4; fbxo8
25 – Signal transduction	9	grin1b; grid1b; pik3r6a; dusp3a; styxl2; map4k2; **bmp4**
26 – Cellular component organisation	9	fyna; col12a1a
27 – Organelle organisation	6	nin; arhgef10; kif20bb
28 – Actin cytoskeleton organisation	8	arhgef10; styxl2; cnn3b; cnn3a
29 – Cellular process	15	cdk5rap3; asah2; pimr23; pimr25; kif20bb; reck; pik3r6a; myo1d; arhgef10; **bmp4**
30 – Process	12	fyna; pik3r6a
31 – Regulation of cellular process	14	reck, adam8b; plaub; rnf213b; sost; dusp3a; styxl2; pik3r6a; **bmp4**; pim2; pimr23; pimr25
32 - Process	7	pck2; stxbp4; bin1b; ank3a

3.6 Discussion

Here we report evidence of segregation of lateral line system phenotypes in hybrids of *A. calliptera* and *A.stuartgranti*, two species with divergent lateral line phenotypes (Figure 3.1) (Webb et al. 2014; Becker et al. 2016; Edgley and Genner 2019) which have evolved from common ancestry within the past one million years (Malinsky et al. 2018). This adds to existing evidence for the heritability of variation in neuromast number and patterning, as has been shown in sticklebacks (Wark et al. 2012) and provides additional evidence for heritability of canal pore architecture in cichlid fishes (Scott et al. 2023; also see Chapter 4).

3.6.1 Phenotypic disparity in hybrids

For all four cranial canals we measured, we found that *A. stuartgranti* had significantly larger pores than *A. calliptera*, and that the pooled F_2 and F_3 hybrid group was intermediate between the two (Figure 3.2). This was also the case for most neuromast count variables we measured, such as the number of trunk canal (Figure 3.3C) and superficial neuromasts (Figure 3.3D). However, for the total number of superficial neuromasts on the head (Figure 3.3B) and the gill region of the head (Figure 3.3C), the range of observed phenotypes of hybrids exceeded the extremes shown by the parental phenotypes: for example, *A. stuartgranti* exhibited a maximum of 76 visible superficial neuromasts on the head, whereas this was up to 116 among hybrids. In addition, the minimum number observed for *A. calliptera* was 29, whereas for the hybrids this was 8. This may be due to limitations in the sample size of our parental populations, meaning they are not representative of the full extent of variation. Alternatively, it may be due to inconsistencies in the staining and imaging process, despite the process being identical for every individual, and neuromasts counts only being accepted where neuromasts were clearly identifiable.

Another explanation is transgressive segregation in hybrids (Rieseberg et al. 1999), causing them to exhibit extreme phenotypes beyond the magnitude of those in the parental lines. This has been noted in multiple morphologies in related cichlid species (Bell and Travis 2005), specifically in craniofacial morphology (Albertson and Kocher, 2005; Holzman and Hulsey, 2017; Husemann et al. 2017) which is closely related to – and may partially constrain evolution of – the cranial canal lateral line system (Edgley and Genner, 2019). Transgressive segregation has been proposed as one of the mechanisms by which hybridisation can promote adaptive radiation (Kagawa and Takimoto, 2018). If transgressive segregation of neuromast number is common in cichlids, this may be one mechanism facilitating niche adaptation during adaptive radiation, through accelerated generation of niche-appropriate mechanoreceptive morphologies.

3.6.2 Genomic loci associated with lateral line system variation

Our approach enabled the construction of a list of candidate genes linked to genomic loci involved in lateral line system disparity in Malawi cichlids. Notably, we find little overlap between genes associated with cephalic canal pore morphology (Figure 3.4A-D) and the genes associated with neuromast count (Figure 3.4E-K). This may indicate that genomic mediation of lateral line phenotypes is highly modular. In addition, as well as varying broadly between cephalic canal and neuromast modules, it also appears to vary within these systems. Each canal of the head we investigated, for example, appears to have a somewhat distinct genomic underpinning (Figure 3.4A-D). This was also observed by previous studies correlating genomic loci with neuromast count in sticklebacks (Wark et al. 2012), although in that case *Eda*, a highly epistatic and broad-effect gene, was widely associated with superficial neuromast variation (Wark et al. 2012; Greenwood et al. 2016; Robertson et al. 2017). To determine the exact genetic mechanism driving disparity in each of these cichlid lateral line phenotypes separately, a higher mapping resolution will be required through increased sample size. This could be augmented by verification of candidate loci, through gene expression analyses or gene editing approaches.

3.6.3 Gene ontology

For all gene enrichment analysis on candidate SNPs from GWAS, and subsequent semantic clustering analysis, we observed proportionately large clusters of neurogenesis and growth-related GO terms (Figure 3.5; Figure 3.6). In the case of the pooled analysis, this formed the largest cluster (Figure 3.5A). Identifying clusters of nervous system related GO terms is compatible with expectations for a study of lateral line phenotypes, and as such these findings validate our approach of using genome-wide association analyses on data for hybrids of an interspecific cross. GO terms enriched in hybrids included endodermal development, axon regeneration, and neuronal development (Figure 3.5B), among others, all of which may point to genes involved in the origin of diversity in lateral line system morphology.

One cluster of 10 GO terms from the neuromast count SNPs was annotated as being specifically related to lateral line system development (Figure 3.6B). When investigating the genes involved with these terms, we found that the *nsfa* gene (previously known as *nsf*; UniProt B7ZV62) was associated with multiple enriched lateral line system GO terms (Table 3.1). The *nsfa* gene has been associated with lateral line nerve development in zebrafish (Philip et al. 2012), and in analyses investigating neurotoxicity (Prats et al. 2017). It is also known to be actively expressed in lateral

line hair cell (kinocilium) tissues, as revealed by RNAseq and transcriptome analyses (McDermott et al. 2007). Specifically, it has been found to be important for maintaining synaptic connections in hair cells of zebrafish lateral line neuromasts (Mo and Nicolson, 2011). We also identified the gene *bmp4* (UniProt F1QAD6) in several enriched GO terms. *Bmp4* has been characterised as an important gene in craniofacial development across many taxa including Darwin's finches (Abzhanov et al. 2004), but also in cichlid fishes (Albertson et al. 2003). It has been found to be related to adaptation to different trophic niches during cichlid radiation, facilitating trophic morphology modifications (Albertson et al. 2003; Streelman and Albertson, 2006; Gunter et al. 2013; Carruthers et al. 2022), including the lower oral jaws, which also contain the ventral cephalic lateral line canals (Edgley and Genner 2019).

3.6.4 Limitations of this study

The commonly used approach when testing for associations between genotype and phenotype in a hybrid interspecific cross design is quantitative trait locus (QTL) mapping (Albertson et al. 2003; Parsons et al. 2012; O'Quin et al. 2013). QTL analyses must be supported by construction of a linkage map (Manichaikul et al. 2010; Sticca et al. 2021), which requires a pedigree of all included individuals. Due to constraints in our experimental design, this was unavailable without genotype imputation of individuals for which data was missing (Davies et al. 2016), and inference of relatedness through identity-by-descent (IBD) allele sharing analyses or similar. As both of these were not feasible, we used the alternative approach of conducting a genome-wide association study on a related datatset. This strategy has precedent in the literature and been used successfully with F_2 populations in the past, as an approach to conducting association mapping (Lutz et al. 2017; Wang et al. 2017).

Despite this, such GWASs on related individuals typically require far larger sample sizes than ours (>1000). Furthermore, a major principle of most GWASs, is that they conventionally require that all individuals be less related than second-degree, as this can lead to biases in the estimation of effect sizes of SNP associations in linear mixed models (Marees et al. 2018). Here, we correct for population structure by including a kinship matrix (Zhou and Stephens, 2012) and the first five components of a genomic PCA as covariates in all GWAS analyses (Zhao et al. 2018). These may serve to partially account for the highly inflated relatedness between individuals in our mapping population. However, in order for the effect of the population structure of highly related individuals, a more rigorous approach would be to infer individual relatedness into pedigrees and conduct QTL mapping.

Further to this, GWASs often require the statistical power provided by sample sizes of close to or exceeding 1,000 individuals, though they have been in the past successfully conducted with far lower numbers of individuals, such as our dataset (Klein 2007). As such, in order to fully validate our results, power analysis must be conducted to determine the sample size required. The sample size requirements of GWAS are dependent on the nature of the phenotype, and how polygenic the trait is likely to be, among other factors. Taken together, our results must be viewed in the context of these constraints, and considered as an initial look into the genomic underpinning of lateral line phenotypes, with further full validation of our findings required through further analyses and expanded sample sizes.

3.6.5 Conclusions

Here, we provide the first insights into the genomic basis for diversity of lateral line system disparity among cichlid fishes. These findings, combined with those from previous research on sticklebacks (Wark et al. 2012; Archambeault et al. 2020), provide a basis upon which further research may be conducted. The results constitute a list of genes to investigate in the context of existing datasets, such as whole-genome sequences of all ~800 species of the Lake Malawi radiation (Malinsky et al. 2018; Svardal et al. 2020), the nearby Lake Victoria and Tanganyika radiations (Meier et al. 2017; Ronco et al. 2021), and datasets involving New World cichlids (Kautt et al. 2020). In addition to this, verification of these results may be provided by gene expression analyses, through RNAseq of focal tissue, including oral jaw bone, lateral line scales or endodermal tissue of the head and trunk. It may also be possible to use CRISPR/Cas9 editing to verify the function of these genes, similarly to previous work on cichlid pigmentation (Kratochwil et al. 2018; Clark et al., 2022).

3.7 Supplementary Information

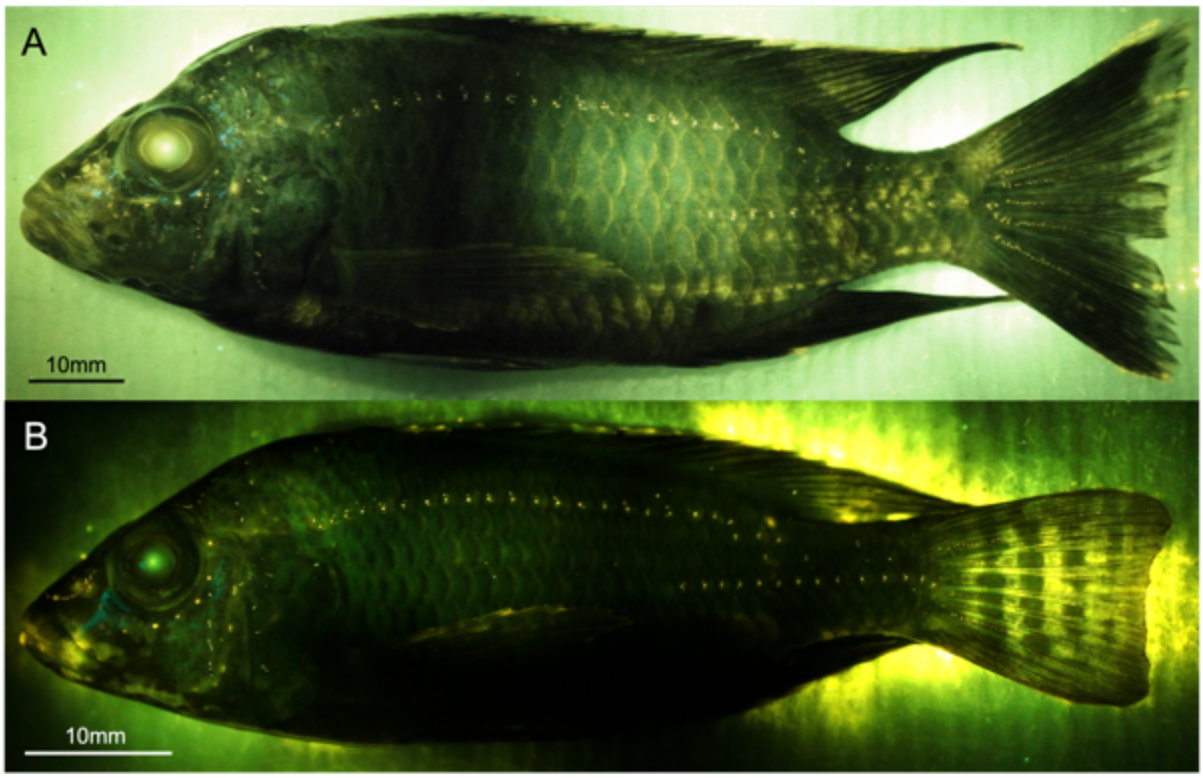


Figure S3.1. The lateral line system of *Aulonocara stuartgranti* and *Astatotilapia calliptera*. (A) Lateral view of a DASPEI stained *A. stuartgranti* highlighting superficial and canal neuromasts. (B) Lateral view of a DASPEI stained *A. calliptera* highlighting superficial and canal neuromasts. Asterisks indicate approximate locations of canal neuromasts. CT scans are annotated with the names of the cranial bones which contain lateral line canals. MD = mandibular canal; IO = infraorbital canal; PR = preopercular canal; SO = supraorbital canal; OT = otic canal; PO = postotic canal; ST = supratemporal canal; TC = trunk canal.

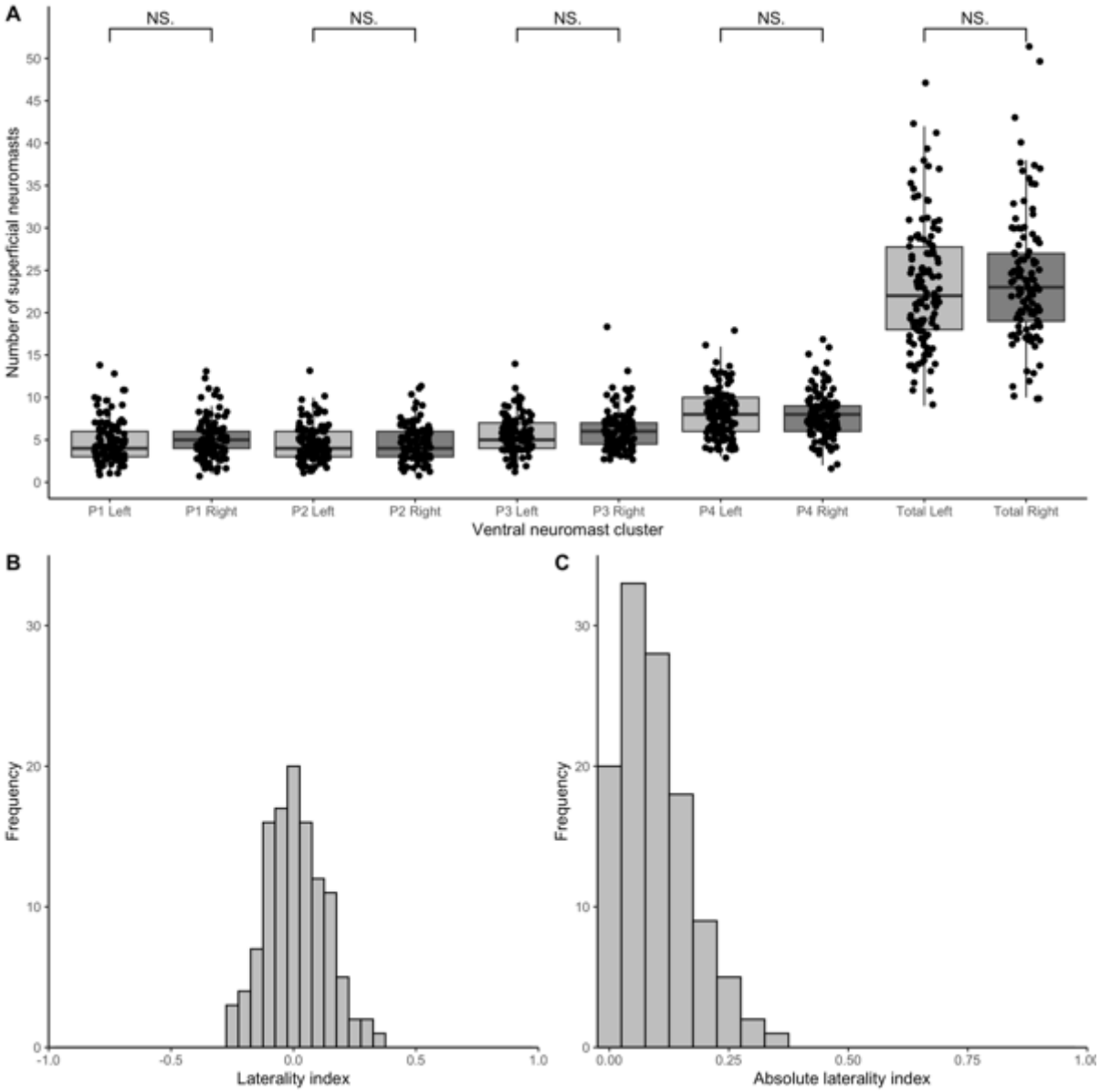


Figure S3.2. No evidence of lateralisation in the numbers of neuromasts of the ventral head in our F_2s. (A) Comparisons of the numbers of superficial neuromasts clustered around each canal neuromast in the lower jaw on corresponding locations of the left and right sides of the body. There is no evidence of any significant differences between the two. (B) Laterality index for the numbers of neuromasts on each side of the body. (C) Absolute laterality index for both sides of the body.

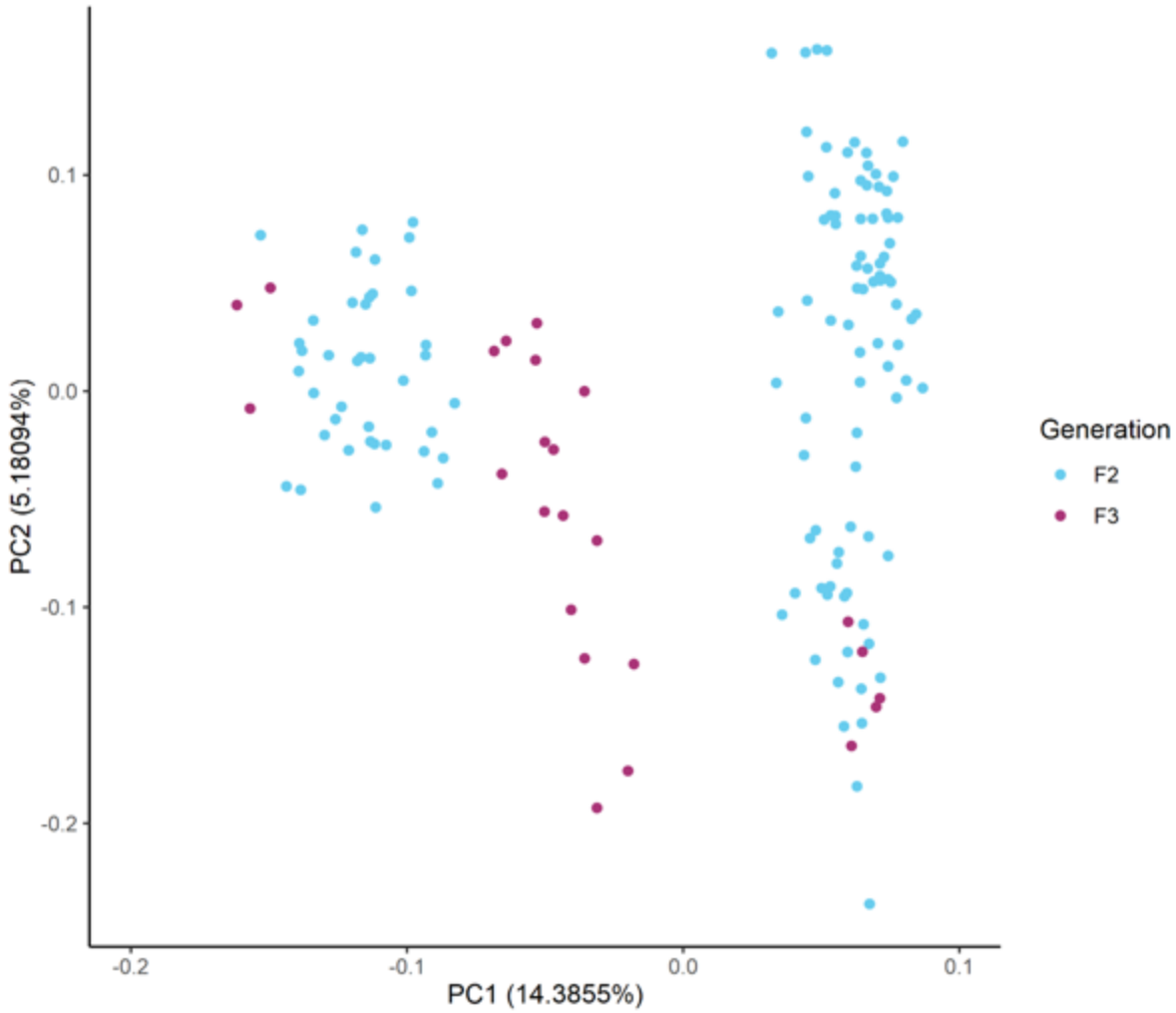
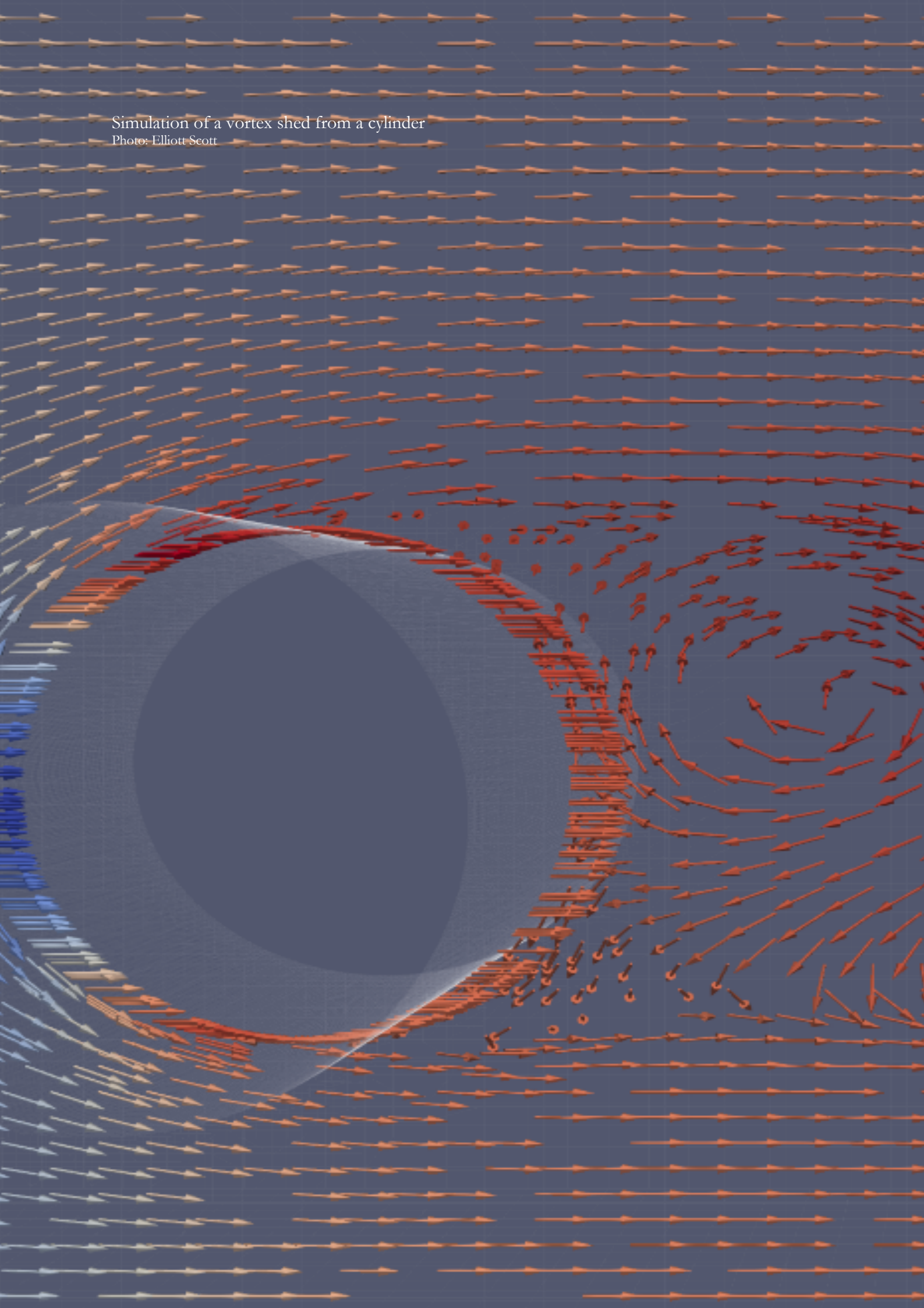


Figure S3.3. Principal components analysis on genomic variation in hybrid individuals of our cross. Here we show clear splitting of F_2 individuals into two groups and F_3 individuals into 3 clusters along the main axis of genomic variation, PC1. This is consistent with the predicted segregation of alleles in an intercross, and validates our decision to include PCs as covariates in each GWAS, given the clear population structure present along both axes shown here.



Simulation of a vortex shed from a cylinder
Photo: Elliott Scott

Chapter 4

The role of the lateral line system in sensory perception and collective behaviour in cichlid fish

4.1 Abstract

The lateral line system of fishes provides cues for collective behaviour, such as shoaling, but it remains unclear how anatomical lateral line variation leads to behavioural differences among species. Here we studied associations between lateral line morphology and collective behaviour using two morphologically divergent species and their second-generation hybrids. We identify collective behaviours associated with variation in canal and superficial lateral line morphology, with closer proximities to neighbouring fish associated with larger canal pore sizes and fewer superficial neuromasts. A mechanistic understanding of the observed associations was provided by hydrodynamic modelling of an artificial lateral line sensor, which showed that canal-based neuromasts were less susceptible to saturation during unidirectional movement than superficial neuromasts, while increasing canal pore size elevated sensitivity to vortices shed by neighbouring fish. Our results provide insight into mechanisms behind lateral line flow sensing during collective behaviour in fishes.

4.2 Author Contributions

This chapter has been published as a research paper in Royal Society Open Science.

This chapter was a collaborative project with a major contribution from both **Duncan Edgley** and **Elliott Scott**. As such, the data included in this chapter were also submitted as part of a Doctoral Thesis at the University of Bristol by **Elliott Scott**.

Duncan Edgley conducted the CT scans, analysed pore morphology, assisted with analysis of neuromast imaging, co-analysed the data and co-wrote the manuscript. **Elliott Scott** performed the behavioural experiments, imaged the neuromasts, performed hydrodynamic modelling/simulations, co-analysed the data and co-wrote the initial manuscript. Study design, initial experimental design was by **Martin Genner**, **Christos Ioannou**, and **Sabine Hauert**, in collaboration with **Duncan Edgley** and **Elliott Scott**. All aforementioned authors analysed the data and wrote the final manuscript. **Thomas Davies** and **Elizabeth Martin-Silverstone** assisted with CT scanning. **Domino Joyce** and **Alan Smith** bred the hybrids in aquaria at Hull University.

4.3 Introduction

The lateral line system, sometimes described as a 'touch-at-a-distance' sense (Dijkgraaf 1963), is used by fishes to detect changes in water flow and pressure and capture information about their surroundings. It is comprised of mechanoreceptors (neuromasts) that are either within subdermal channels (canal neuromasts) or on the surface of the skin (superficial neuromasts) (Coombs et al., 1988; Beckmann et al., 2010). Superficial neuromasts can be considered as velocity detectors, used to detect the direction and velocity of flow, while the canal neuromasts may be considered as pressure gradient detectors, sensing differences in water movement between adjacent canal pores, primarily caused by turbulence (Kroese and Schellart, 1992; Mogdans 2019). Neuromasts of the lateral line system are distributed over the head, trunk and tail of the fish, and species can differ considerably in the size and position of lateral line components (Kroese and Schellart, 1992; Hoekstra and Janssen, 1986).

The lateral line system plays an important role in a number of behaviours, including prey detection (Hoekstra and Janssen, 1986, Janssen et al. 1995), alignment to flow (rheotaxis) (Montgomery et al. 1997; Baker and Montgomery, 1999; Simmons et al. 2004; Olszewski et al. 2012; Suli et al. 2012), mate selection (Medina et al. 2013; TerMarsch and Ward, 2020), conspecific aggression (Butler and Maruska, 2016; TerMarsch and Ward, 2020) and shoaling (Pitcher et al. 1976; Partridge and Pitcher, 1980). It has particular importance in situations where other senses are limited, for example by enabling detection of prey in dark (Janssen et al. 1995; Schwalbe et al. 2012; Schwalbe and Webb, 2014) or turbid (Rowe et al. 2003; Chamberlain and Ioannou, 2019) conditions, or detecting prey buried within substrate (Schwalbe et al. 2016). Fish can use canal neuromasts to sense vibrating obstacles (Blaxter and Batty, 1985; Janssen and Corcoran, 1998) and neighbouring fish, and are able to use their lateral line systems to shoal (Pitcher et al. 1976; Partridge and Pitcher, 1980). When the sensory input from the lateral line system is interrupted through surgical methods or chemical treatment, multiple behaviours are either severely impaired or lost, including normal shoaling ability (Pitcher et al. 1976; Partridge and Pitcher, 1980; Faucher et al. 2010), foraging (Schwalbe et al. 2012; Schwalbe and Webb, 2014; Schwalbe et al. 2016), competitor assessment (Butler and Maruska, 2016; TerMarsch and Ward, 2020) and mate selection (Medina et al. 2013; TerMarsch and Ward, 2020).

Typically, studies of lateral line function have compared control fish to those in which lateral line function has been chemically or physically impaired. While these treatments are effective at inactivating the lateral line (e.g. Song et al. 1995), they can cause other side-effects. Aminoglycoside

antibiotics and cobalt chloride damage auditory hair cells (van Trump et al. 2010; Brown et al. 2011) and olfactory receptor neurons (Janssen 2000; Butler et al. 2016), respectively. Cobalt chloride is also toxic to fish even at low exposure (Janssen 2000; Butler et al. 2016) and even what is considered a safe dose for lateral line ablation has a 15% mortality rate in zebrafish (Stewart et al. 2017). Additionally, some studies have claimed full or partial lateral line deactivation (Pitcher et al. 1976; Partridge and Pitcher, 1980; Song et al. 1995), where this may not have been the case (Pitcher et al. 1976; Partridge and Pitcher; van Trump et al. 2010; Brown et al. 2011). Lateral line transection is another method sometimes used for ablation (New et al. 2001; Butler and Maruska, 2016), but this requires invasive surgery and effects on fish health have not been studied. Thus, based on previous work, it has been difficult to confidently determine the exact function of the lateral line system or its various components, in relation to one another or other sensory systems.

Here, we study associations between morphological and behavioural variation in second generation (F_2) hybrids of two phenotypically divergent Lake Malawi cichlid fishes, *Aulonocara stuartgranti* and *Otopharynx lithobates* (Edgley and Genner, 2019) (Figure 4.1). These species are part of the Lake Malawi cichlid adaptive radiation, in which hundreds of ecomorphologically-divergent species have evolved from common ancestry in the last million years (Malinsky et al. 2018; Svardal et al. 2020). These species possess extensive interspecific variation in lateral line systems (Malinsky and Salzburger, 2016; Edgley and Genner, 2019). *Aulonocara stuartgranti* is widely recognized as a representative of species with a 'wide' canal morphology (Schwalbe et al. 2012; Schwalbe and Webb, 2014; Konings 2016; Schwalbe et al. 2016), while *Otopharynx lithobates* is typical of species with a 'narrow' canal morphology, which are more commonplace in the Lake Malawi radiation (Greenwood and Kornfield, 1983; Konings 2016). Expanded 'wide' cranial canal pores have been associated with improved ability to catch prey in front (New et al. 2001) and above fish (Saunders and Montgomery, 1985; Bleckmann 1988) and even hidden in the substrate (Schwalbe et al. 2012; Schwalbe and Webb, 2014; Schwalbe et al. 2016). Thus, diet seems to have a strong association with canal pore variation between cichlid species (Edgley and Genner, 2019). Equally, since superficial neuromasts in several regions of the anterior lateral line system have been associated with shoaling tendency (Greenwood et al. 2013), we hypothesised that superficial neuromasts on the head and body may take a significant role in collective behaviour in cichlids. In this study, we use the term collective behaviour to refer to shoaling or schooling behaviour, given that reliable alignment information is lacking.

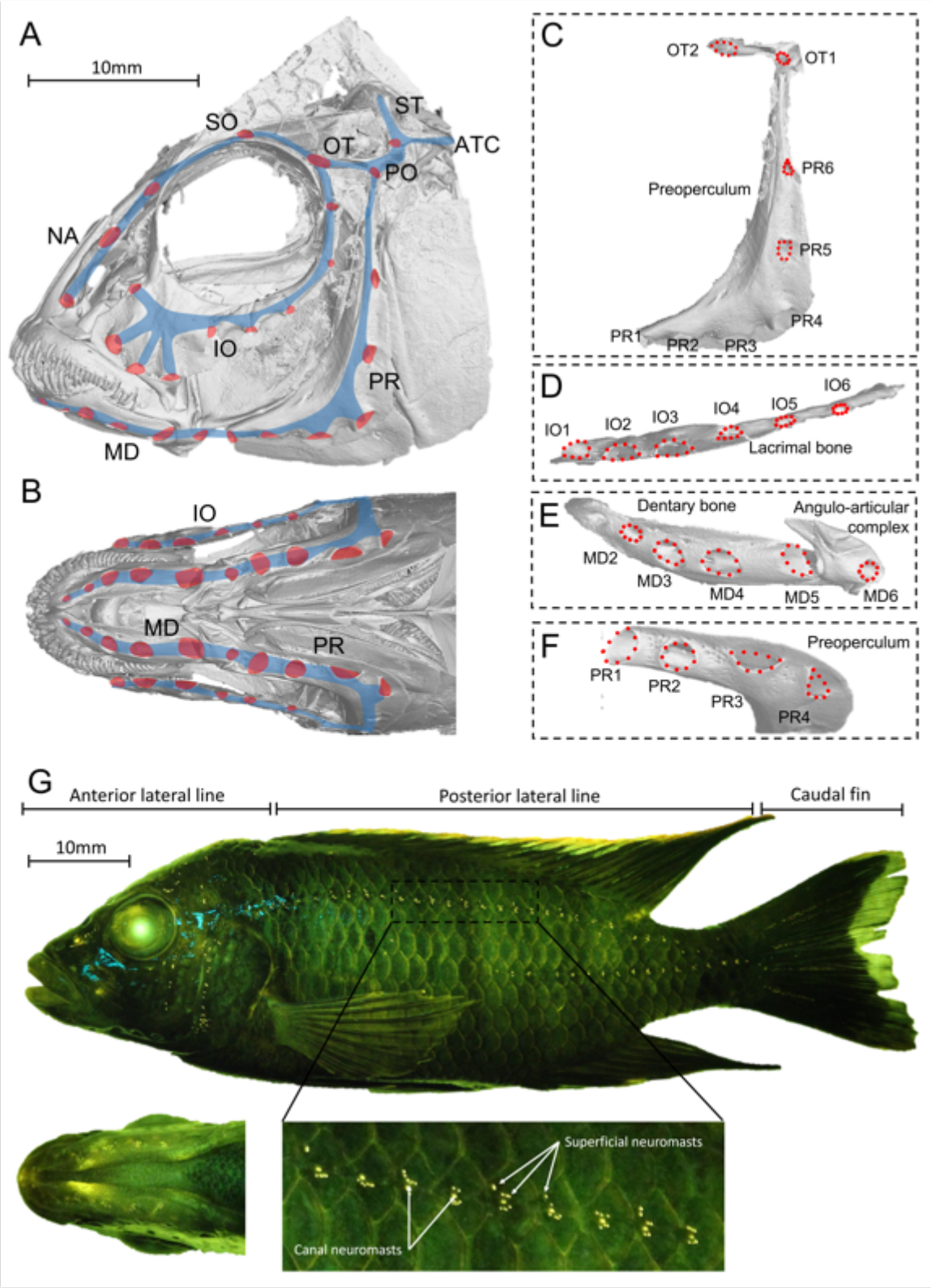


Figure 4.1 The lateral line system morphology quantified in this study, illustrated on an example F2 hybrid of *Aulonocara stuartgranti* and *Otopharynx lithobates*. (A) Lateral view CT scans showing anterior canal lateral lines and pore structure. (B)Ventral view CT scans showing anterior canal lateral line and pore structure. (C) Expanded lateral view of preopercular and otic canals with pores marked by dotted red circles. (D) Expanded lateral view of infraorbital canal with pores marked by dotted red circles. (E) Expanded ventral view of preopercular canal with pores marked by dotted red circles. (F)

Expanded ventral view of mandibular canal with pores marked by dotted red circles. (G) DASPEI stained superficial and canal neuromasts, showing the definitions of regions used in this study. Shown is the lateral view of the body, and the ventral view of the head. In addition, we show an expanded view of DASPEI stained canal and superficial neuromasts in a typical section of posterior canal. IO = infraorbital; MD = mandibular; OT = otic; PR = preopercular; NA = naris; SO = supraorbtial; PO = post-otic; ST = supratemporal; ATC = anterior trunk canal.

We also present the results of a hydrodynamic modelling study, investigating neuromast responses under a variety of different conditions. We specifically explored the concept that superficial neuromasts reach a state of saturation, or maximum deflection, due to background flow, which could prevent them from detecting certain flow stimuli. In contrast, the structure of the canals allows them to detect these otherwise masked flow changes, like those caused by vortices shed from upstream companions by filtering hydrodynamic noise (Figure 4.2A-B) (Dijkgraaf 1963; Kroese and Schellart, 1992; Mogdans 2019). We also simulate a novel design of artificial lateral line and demonstrate its ability to detect the vortices shed by an upstream cylinder as an approximation of the wake of a swimming fish (Figure 4.2C) (Liao 2007; Fish 2010).

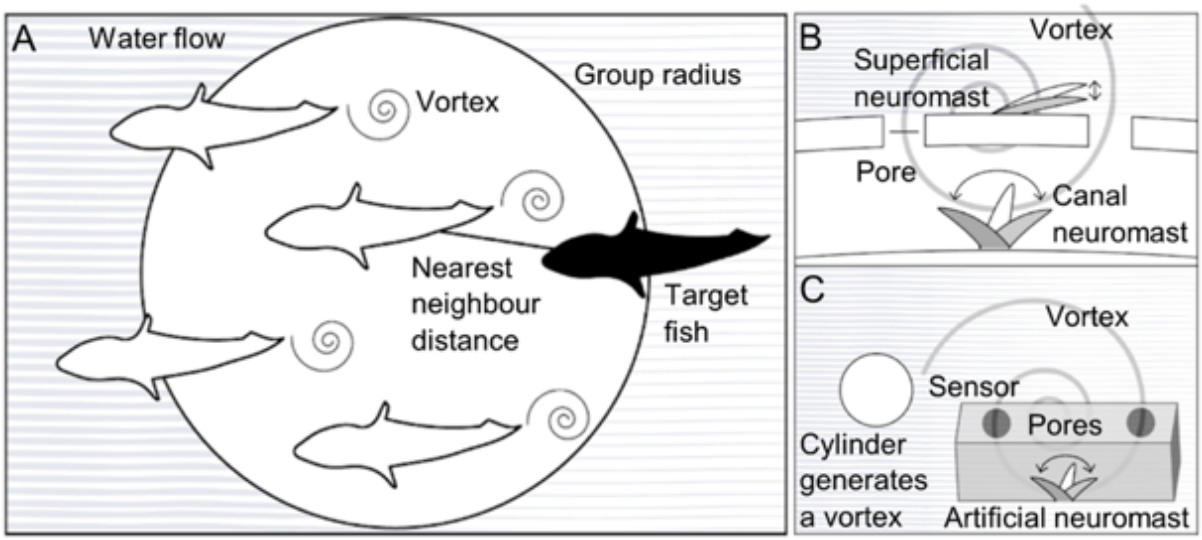


Figure 4.2 Schematic diagram of our study elements. (A) A group of fish shedding vortices, as a summary of the response variables included in our behavioural experiments' models: nearest-neighbour distance and group radius. Parallel lines are used to indicate background laminar flow. (B) Response of superficial and canal neuromasts to a vortex when in presence of background flow. The superficial neuromast is more affected by background flow, whereas the canal neuromast is less affected. The canal structure filters out background laminar flow (pale grey lines) while allowing turbulent flows like shed vortices to be detected more clearly. A stiffer superficial neuromast would be less affected by background flow, but this then has the negative side effect of causing smaller deflections in response to the important flow features, such as shed vortices. (C) A cylinder can be used to generate vortices (approximate to those generated by an upstream fish), and an artificial canal structure containing an artificial neuromast can detect them.

Overall, this study aims to investigate how differences in lateral line morphology, particularly differences between the two lateral line submodalities, cause changes to collective behaviour with sensory ability being the presumed driver. Observations about how sensing ability appears to affect collective behaviour are then explored in simulation to verify predictions. The work sheds new light on the role of the lateral line in collective behaviour of fishes and takes steps towards a novel artificial lateral line design.

4.4 Materials and Methods

4.4.1 Cross design and sample collection

In total, 116 focal fish were included in our analysis: ten *Aulonocara stuartgranti*, ten *Otopharynx lithobates*, and 96 *Aulonocara stuartgranti* × *Otopharynx lithobates* second-generation (F_2) hybrids from an intercross (Table S4.1). Our hybrids had a mean total length of 117.31mm (σ = 17.71), with *A. stuartgranti* measuring a mean total length of 110.50mm (σ = 10.28), and the *O. lithobates* measuring a mean total length of 117.10 (σ = 13.41). Second-generation *A. stuartgranti* × *O. lithobates* hybrids exhibit substantial variation in their cranial lateral line systems, between the 'wide' and 'narrow' canal morphologies of their parents, and collectively these hybrids span much of the range of variation seen across the Lake Malawi radiation (Figure 4.1) (Webb 2014). Morphological data from the Lake Malawi radiation is used to frame the results of this study, giving better context and demonstrating that the two parent species are typical examples of the 'wide' and 'narrow' morphologies; see Edgley and Genner (2019) and Chapter 2 for more in-depth analyses.

The use of F_2 hybrids allows for decoupling of any genetically unlinked lateral line traits, as well as other sensory traits that enable the perception of environmental cues, such as vision and olfaction. Both parent species, along with many other species of cichlid, including our companion species *Hemitilapia oxyrhynchus* (a common species from the Lake Malawi radiation) exhibit collective movement and shoal in natural populations (Oliver 1984; Robinson and Turner, 1990; van Dongen et al. 2014; Moss et al. 2015; Konings 2016; Thünken et al. 2016; Marlin et al. 2019; Durrer et al. 2020), making them a useful candidate for these collective behaviour experiments. Behavioural analysis showed that these companion fish tended to group tightly during experiments (figure S4.1). While cichlids in Lake Malawi, being lacustrine, are not usually exposed to constant flows, the low velocity flows here enable us to more reliably encourage swimming motions (in the absence of flow, fish showed an increased tendency to remain further apart; figure S4.5).

Cichlids were bred in aquaria at the University of Hull. Ten individuals from each of the parent species were housed separately in 90 L tanks and approximately 100 hybrid individuals were housed in a 720 L tank. A further 90 L tank held approximately 30 individuals of the species used as companion fish, *Hemitilapia oxyrhynchus*. All tanks were kept on a recirculating water system at 23°C and a 12-12 light-dark cycle. They were fed a varied diet, consisting of ZM large granular pellets, tetra tropical flake, frozen blood worm, frozen prawn, frozen brine shrimp (+ brine shrimp with supplements such as spirulina, garlic, aloe), vegetarian diet blister packs, cichlid diet blister packs, mysids, krill, daphnia, white mosquito larvae and tubifex.

4.4.2 Behavioural experiments

In this study, we attempt to quantify collective behaviour. This term is used over shoaling or schooling due to a lack of reliable data on the orientation of individuals to calculate variables such as polarisation. During trials, focal fish were filmed in an experimental flow tank, in a group with four companion fish, in both laminar and turbulent flows (figure S4.1). The experimental tank was an arena of 2 × 1 m, within which foamed PVC walls were added to mimic a convergent nozzle to help control flow (figure S4.1). A baffle was added upstream to allow flow to laminarise, in addition to meshes (6mm square holes separated by 3 mm walls). The area measured 121 × 55.78 × 20 cm. For the turbulent treatment, turbulence was generated by vertical 32 mm clear plastic rods attached to the downstream side of the flow straightener (Figure S4.1). The experimental tank was white and brightly lit to give fish ideal visual conditions, so that any differences in visual acuity would be less pronounced and be less likely to affect behaviour.

Each trial consisted of one individual of either the hybrid or parent species (focal individual) and four individuals of the companion species. This was to reduce between-trial variation, based on morphological and behavioural differences between individuals of the companion species. With a decreasing number of companion individuals, differences in shoaling behaviour between trials would be increasingly affected by variation between the companion fish, rather than variation between the focal individuals. By varying the morphology of only one fish (our focal individual) and keeping the traits of the other group members approximately constant, variation not accounted for by the morphology of the focal individual is reduced. The five fish in each trial were given ten minutes to acclimatize with no flow in the experimental arena. As the morphological measurements were conducted after the behavioural experiments, fish were caught blind regarding lateral line system morphology. Trials lasted for 20 minutes at an average flow speed of 7cm/s. This was calculated by measuring the distance that particles flowed downstream at several

positions in the tank and then taking the average. This was determined to be high enough to encourage swimming and rheotactic behaviours, but not so high as to exceed flows they would experience in their natural environment (Oliver 1984; Konings 2016; Malinsky and Salzburger, 2016). Following each trial, the flow was switched off again and fish were given a further 10-minute rest period. During these 10 minutes, turbulence generators were either removed or inserted, depending on which treatment (turbulent or laminar) was given first. The flow was then switched back on for the second 20-minute trial.

Each test shoal was thus tested in both turbulent and laminar treatments, with half experiencing laminar first and half experiencing turbulent first (randomly assigned); the same companion fish were used in both the laminar and turbulent trial for a given target individual. The overall pool size of companion fish was kept small to lessen the effects that variation in this population might have on the collective behaviour metrics. 28 companion fish were used in total. It was expected that the large number of trials and the constant mixing of companion individuals would average out any effects of individual variation between companion fish over the experiment. After the testing, the focal individual (i.e. hybrid or parent) was placed in isolation in a smaller 45 L tank to await further steps of the experiment, while the companion individuals were kept separate for the remainder of the day to prevent them being tested again. All procedures were approved by the University of Bristol Ethical Review Group (UIN/17/005).

4.4.3 Video recording and analysis

Recordings were made using a Panasonic VX870 camera, filming in 4K at 25fps. The camera was suspended 148 cm above the base of the experimental tank. We zoomed the camera such that the experimental arena filled the entirety of the screen. The camera filmed in nine minute segments which were stitched together to form a single 20 minute video per trial. Quality was reduced to 1080p using ShotCut (https://shotcut.org/). The automated video tracking software idTracker version 2.1 (Pérez-Escudero et al. 2014) was used to extract x and y positional coordinates, which were analysed in MATLAB R2013b.

To determine the focal fish's nearest neighbour, we determined the distance between all individuals in every frame of video. The lowest of these pairwise values for the focal individual was considered the nearest neighbour, and mean across all frames was calculated from this. Group radius was calculated by determining the central point of the shoal (the mean of all individuals' coordinate points) in each frame, then calculating the distance to the individual furthest from the centroid

(Figure 4.2A). This was highly correlated with another alternative measure of group cohesion, the bounding hull circumference (Figure S4.3; Figure S4.5). We also calculated the focal fish's distance to the group centroid, which was also found to be highly correlated with group radius (Figure S4.3). Nearest upstream neighbour and nearest 'field of flow sensing' neighbour metrics were calculated from the nearest neighbour distances (Figure 4.2). Neighbours were deemed upstream if they had an x coordinate value lower than that of the target individual, while those deemed in the 'field of flow sensing' had both a lower x-coordinate, and a y-coordinate within 5 cm of the target individual. The origin of the coordinate system was at the most upstream and furthest left point of the experimental tank, and the environment was oriented such the x-axis pointed downstream, and the y-axis pointed cross-stream. Values where neighbours were not in front, or not directly in front respectively, were excluded. These metrics were included because it is not possible to detect downstream individuals using the lateral line since hydrodynamic stimuli generated by tail beats cannot travel upstream. In this way, we attempt to better ensure that observed behaviours are informed by lateral line system input. We watched each video in idPlayer (Pérez-Escudero et al. 2014) to verify that the identity of the focal fish remained the same within the two 20-minute trials.

4.4.4 DASPEI staining

After three focal fish were tested, focal individuals were transferred to a 0.01 mg/ml DASPEI solution for 30 minutes. They were then euthanised by submersion in a 300 mg/L solution of tricaine methanesulfonate (MS-222) for ten minutes and confirmation of death by destruction of the brain, according to an approved schedule one method. We then imaged each individual beneath a Sigma 18-200 mm f/3.5-6.3 lens on a Canon 550D DSLR camera. We used a Royal Blue lamp (465nm) to excite the DASPEI stain, and a yellow long-pass filter (500 nm) lens to remove interference from the excitation light and allow the emission light to be seen more clearly. Multiple stills were taken of each individual at a resolution of 5184 × 3456 with the lamp redirected to highlight different areas of the fish to ensure that all DASPEI tags were visible. These images were then compiled using the ImageJ (Schneider et al. 2012) stack feature to generate a single image showing all areas of the fish highlighted. The result was a lateral view of the whole fish. Images captured in this way gave sufficient detail to count the canal and superficial neuromasts visible at the surface, henceforth referred to as 'visible superficial neuromasts' or 'visible canal neuromasts'. From this, we quantified six neuromast count variables: anterior superficial, posterior superficial, posterior canal, lower posterior superficial, lower posterior canal and tail neuromasts (Figure 4.1G).

4.4.5 microCT scanning and morphometrics

All individuals were μCT scanned using a Nikon XTH225ST Computed Tomography system in January 2019. We scanned fish two at a time, each scan using 3141 projections and a voxel size of 20-30 μm. μCT parameters were determined following preliminary scans of similar specimens, and careful inspection of the cranial pores following reconstruction. Stacks of images were imported into VG Studio 3.0 (Volume Graphics GmbH) and reconstructed into a 3D model. From these 3D reconstructions, we captured 2D images from the ventral head perspective and lateral head perspective (on the left side of the fish). From these images, four anterior lateral line canal pore size variables were quantified: mean otic canal pore size, mean preopercular canal pore size, mean mandibular canal pore size and mean infraorbital canal pore size (Figure 4.1A-F).

To quantify differences in cranial lateral line canal pore morphology, we used a landmark-based morphometric approach. Using the 2D images generated from CT scanning, we digitized landmarks and semi-landmarks to capture morphological variation using tpsDig 2.30 for Windows (Rohlf 2008). We drew a curve around the circumference of each pore, beginning and ending at a homologous anchor point (the anterior limit of the lateral line canal pore). These curves were resampled resulting in ten equidistant semilandmarks, which were subsequently converted to landmarks in tpsUtil64 1.74 for analysis (Rohlf 2008).

For the images showing the ventral side of the head (Figure 4.1B), we collected landmarks for pores 2-5 of the mandibular canal (Figure 4.1E), pores 1-4 of the preopercular canal (Figure 4.1C), and pores 1-6 of the infraorbital canal (Figure 4.1D) on both sides in each case. From lateral head images (Figure 4.1A), we collected landmarks for pores 5-7 of the preopercular canal (Figure 4.1F) and the two pores of the otic canal (Figure 4.1C). We also measured head size for each specimen, defined as the distance from the anterior limit of the dentary bone to the posterior limit of the operculum. All landmarks were digitized by the same individual, and with only short breaks between digitizing sessions to ensure consistency. We tested for digitisation error by subjecting 20 of our individuals to repeated landmarking. Using Analysis of Variance (ANOVA), we tested for significant differences in mean dentary canal pore area, among both individual and landmarking events (repeat or original landmarking). We found no significant difference between landmarking events (original or repeat) ($F_{(1,19)}$=0.3655, p=0.5526), but highly significant differences between individual specimens ($F_{(19,19)}$=159.77, p<0.001). We visualized this difference by performing a

Procrustes fit on landmark data using R 3.6.1 (R Core Team, 2020), the package geomorph (Adams et al. 2020) and a principal component analysis using procrustes coordinates (Figure S4.2).

Following conversion to landmarks, for each pore, we discounted the tenth landmark (as its coordinates were the same as the first in each instance). We used the remaining nine landmarks' coordinates as vertices to create a polygon, and calculated the area of this polygon using the polyarea function in the 'geometry' R package (Roussel et al. 2019). These areas were used as approximations for pore area for each specimen. Despite these being slight underestimates of the true pore area, our methodology was consistent across all individuals.

4.4.6 Statistical analyses

We used R version 3.6.1 (R Core Team 2020) for all statistical analyses. Generalized Linear Mixed Models with negative binomial distributions were produced using the glmmadmb package (Fournier et al. 2012; Skaug et al. 2013). The response variable in our models was the mean nearest neighbour distance, the mean upstream nearest neighbour distance, the mean 'field of flow detection' nearest neighbour distance or the mean shoal radius. The primary predictor variable of interest was the focal lateral line variable, while species (*Aulonocara stuartgranti*, *Otopharynx lithobates* or hybrid), focal fish's body length, treatment (laminar or turbulent) and trial order (whether it was the first or second trial for this focal fish) were used as our explanatory variables. The ID of the fish was included as the random effect throughout as each focal fish was tested in both laminar and turbulent flow.

To determine which lateral line system explanatory variables explained the most variation in the behavioural response variables, we used the corrected Akaike information criterion (AICc). Specifically, comparisons between AICc value for each model were made using the ICtab function in the bbmle package (Bolker 2017). Comparing models in this way does not give absolute information about each model's predictive power, but a relative comparison of likelihood between models. We thus included in the model comparisons a null model: a model that relates the response variable to all of the explanatory variables, but excludes the primary predictor variable, which was a lateral line variable. Null models included only body length. Models with lower AICc values have higher support and are more likely given the data. We considered a model with an AICc of two or more units lower than the null model (i.e. $\Delta AICc \geq 2$) to have strong support. By comparing AICc values in this way, we determined which aspects of our focal fish's sensory morphology best predicts the shoaling metrics. This method of comparing GLMMs using the AICc was chosen

because mixed effects models are useful when comparing data with more than one source of random variability, in this case, as a result of the repeated measures of position data (five different individuals) over time.

4.4.7 Artificial sensor simulations

We produced a computer-aided design for a model of a cylinder and an artificial neuromast sensor using Autodesk Fusion version 2.0.8816. We chose a cylinder diameter of 100 mm, approximately equal to mean fish body length of our focal fish. The artificial lateral line sensor was designed to mimic a limited section of canal, measuring 25 × 25 × 50 mm, with a highly elastic haircell acting as the neuromast. Our artificial sensor was deliberately macro-scale for twofold reasons: firstly, we wanted to determine whether the behaviour of this canal section at this size can be considered similar to the biologically accurate size; and secondly, to improve the ease with which it may be mass produced. Two evenly spaced holes were cut out of one face to act as canal pores. Five different interchangeable pore diameters were used to investigate the effect that this would have on the artificial neuromast's response to external flow stimuli. This was determined by measuring the mean and range of flow velocities within the sensor, which arise due to external pressure differentials across the pores. A higher range of velocities detected within the sensor indicates a greater response of the neuromast to a stimulus (stimuli were of the same magnitude for all sensors we tested). A rising mean velocity within the sensor indicates a reduction of filtering effect. Flow speed was also measured in the absence of a canal structure, mimicking a superficial neuromast. The sensor was centred on the origin, with the cylinder placed 200 mm upstream direction and 40 mm to one side. We generated .stl files from these models. OpenFOAM (blueCFD-Core 2017) was used to simulate flow speeds of 0.5m/s over 400 seconds. Snappyhexmesh was used to convert the .stl files to be used in the OpenFOAM simulation (Stewart, 2019). MATLAB 2018b and R version 3.6.1 (R Core Team 2020) were used to determine the point of maximum variation, and to generate graphs and plots of flow velocities. Velocity data at the position of the artificial neuromast hair cell was also extracted from OpenFOAM using the 'Plot Over Line' function. This function takes the data from all the mesh points along a desired 1D line, in this case a line drawn perpendicular to the internal face opposite the pores.

4.5 Results

4.5.1 Lateral line system morphology and behaviour of focal fishes

Across the full data set, lateral line morphology and body size were strongly correlated, with larger fishes having increased numbers of visible superficial neuromasts and larger cranial canal pores

(Figure S4.4). After correcting for fish total length, individuals of *A. stuartgranti* were characterized by larger cranial canal pores than *O. lithobates*, but with fewer visible anterior superficial neuromasts (Figure 4.3; Table S1). Hybrids possessed a broad range of phenotypes, encompassing and spanning the two parental phenotypes (Figure 4.3; Figure S4.4; Figure S4.5; Table S4.1).

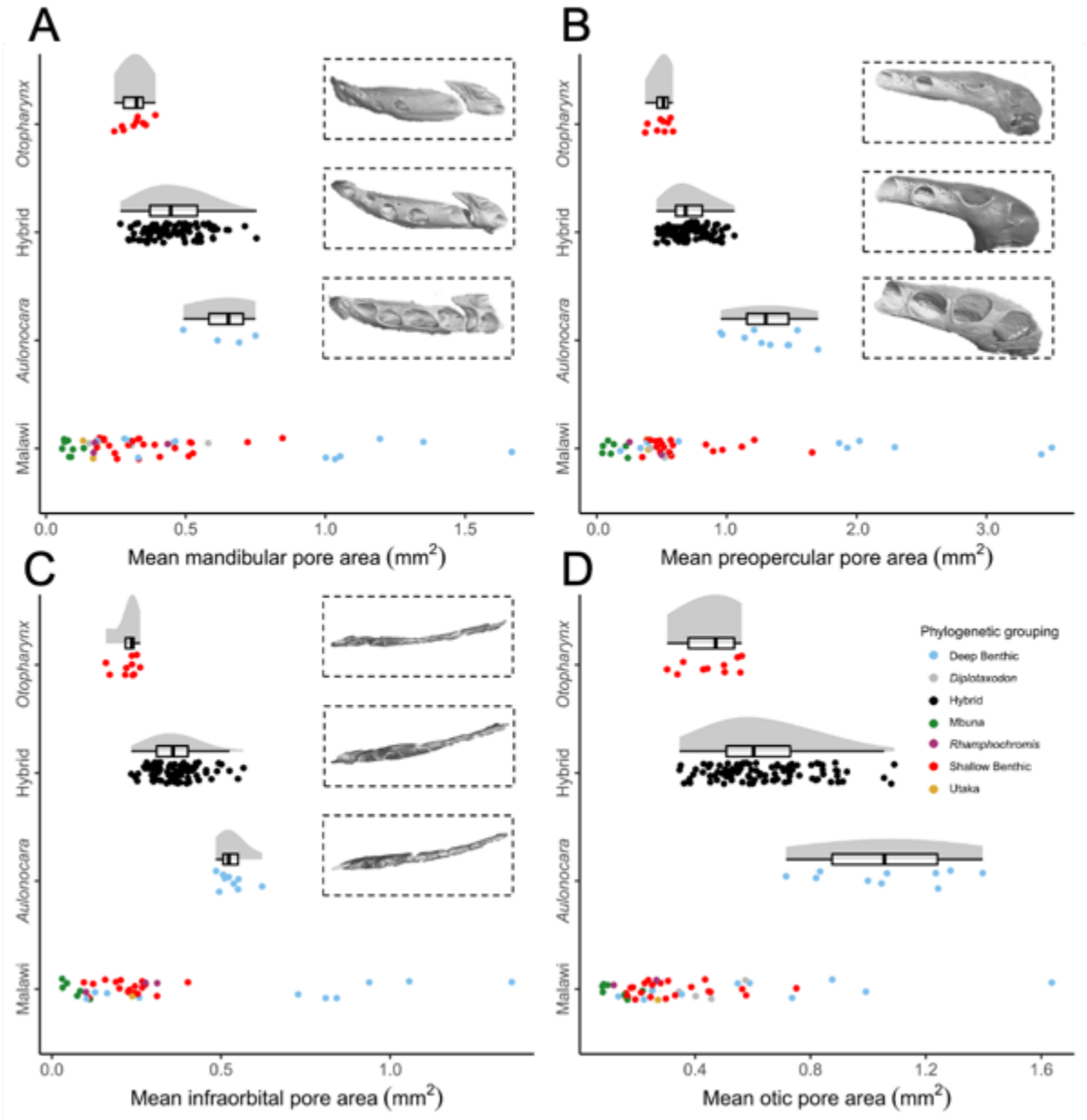


Figure 4.3 Comparative cephalic canal pore morphology of the 116 study individuals (96 F2 hybrids, 10 *Aulonocara stuartgranti*, and 10 *Otopharynx lithobates*). (A) mean mandibular pore areas. (B) mean preopercular pore areas. (C) mean infraorbital pore areas. (D) mean otic pore areas. Mean pore areas of our study individuals are shown in contrast to a mean pore area value for 53 other species of Lake Malawi cichlids (species listed in Edgley and Genner, 2019), from six key phylogenetic groups indicated by colours (electronic supplementary material table S3). Filled black circles represent hybrids, that do not fall into any of the phylogenetic groups. The data for the 53 other species has been taken from a study by Edgley and Genner that analysed the adaptive diversification of the cichlid family (Edgley and Genner, 2019). Additional images of representative CT scans of the different regions of the canal system are also included, with each set of scans corresponding to the graph they are located in.

In terms of behavioural metrics, mean nearest neighbour distance and mean shoal radius of hybrids was intermediate between that of the two parental species (Table S4.1). This pattern persisted when considering only those periods of time when nearest neighbours were either immediately upstream or within the 'field of flow detection' of the focal fish (defined as a channel in front of the fish with boundaries at 50mm either side; Table S4.1).

4.5.2 Associations between lateral line system morphology and behaviour

Increased numbers of both anterior and posterior superficial neuromasts were positively associated with group radii (Figure 4.4A-D; Table 4.1). A higher number of anterior superficial neuromasts was associated with a greater mean distance to nearest upstream neighbour, and neighbour in 'field of flow detection' for the complete dataset. An increased number of posterior superficial neuromasts was also associated with both of the aforementioned collective behavioural metrics, in the hybrids-only dataset.

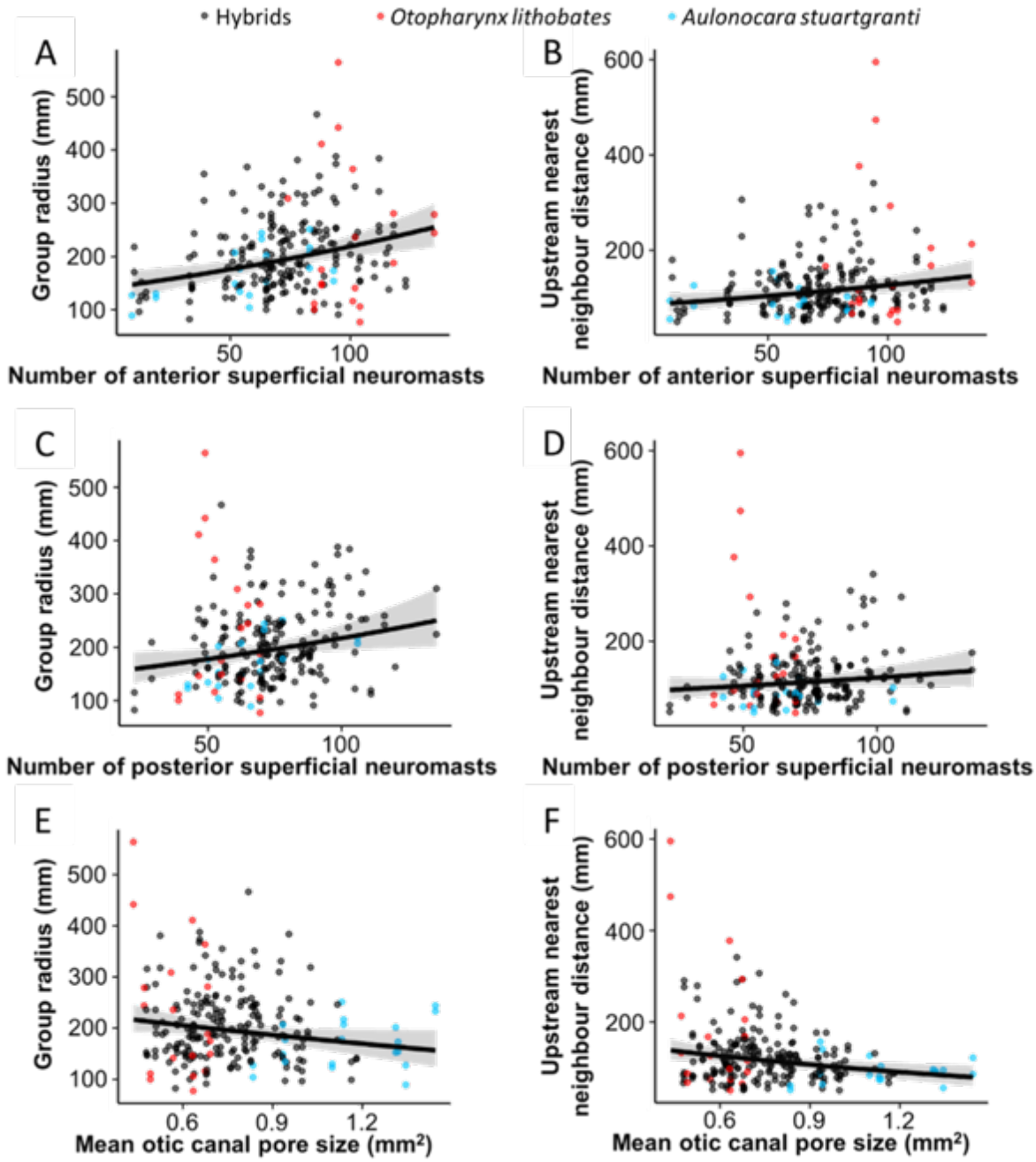


Figure 4.4 Generalized Linear Mixed Models (GLMMs) of associations between lateral line morphology and behaviour. For each plot n=232, including 2 points for each fish, one in the turbulent treatment and one in the laminar. Black points indicate hybrid individuals, red points are *Otopharynx lithobates*, and blue points are *Aulonocara stuartgranti*. (A) Anterior superficial neuromasts and group radius. (B) Number of anterior superficial neuromasts and distance to nearest upstream neighbour. (C) Number of posterior superficial neuromasts and group radius. (D) Number of posterior superficial neuromasts and distance to nearest upstream neighbour. (E) Mean areas of preopercular canal pores and group radius. (F) Mean areas of preopercular canal pores and distance to nearest upstream neighbour. Each comparison here shows a significant relationship, with the exception of the posterior superficial neuromasts and upstream nearest neighbour distance: this was only significant when affecting the hybrids only. Significant relationships are those identified in the AIC comparison process where a model is 2 or more fewer than the null model, which represents a random relationship.

Table 4.1 Summary of statistical models with strong support (ΔAICc ≥ 2). '+' indicates a positive slope between behavioural and morphological variable (green shading), '-' indicates a negative slope (blue shading), U indicates an unsupported result, * indicates a strongly supported interaction between the morphological variable and the treatment (laminar/turbulent flow).

Morphological variable	**Mean nearest neighbour distance**		**Mean group radius**		**Mean nearest upstream neighbour distance**		**Mean nearest neighbour distance within 'field of flow detection'**	
	All	Hybrids	All	Hybrids	All	Hybrids	All	Hybrids
Anterior superficial neuromasts	U	U	+ *	+	+ *	U	+ *	U
Posterior canal neuromasts	U	U	U	U	U	U	U	U
Posterior superficial neuromasts	U	U	U	U	U	+	U	+
Lower posterior canal neuromasts	U	U	U	U	U	U	U	U
Lower posterior superficial neuromasts	U	U	U	U	U	U	U	U
Tail neuromasts	U	U	U	U	U	U	U	U
Otic canal pore size	- *	U	U	U	- *	U	- *	U
Preopercular canal pore size	-	U	-	U	- *	U	- *	U
Dentary canal pore size	-	U	U	U	-	U	- *	U
Infraorbital canal pore size	- *	U	U	U	-	U	- *	U

We identified that there was an interaction with flow regime for each of anterior superficial neuromast models, such that fish with increased numbers of these neuromasts tended to have larger group radii in turbulent flow than in laminar, and tended to be further from upstream neighbours and neighbours in the 'field of flow detection'. For the hybrids-only data set, individuals with more posterior superficial neuromasts formed tighter groups in turbulent regimes than in laminar for both upstream nearest neighbour distance and distance to nearest neighbour in 'field of flow detection'.

Our GLMMs revealed that all measured canal pore variables (mean otic, preopercular, dentary and orbital pore sizes) were negatively associated with nearest neighbour distances when using the complete data set (Figure 4.4E-F). However, for the hybrids-only data, we found no supported associations between canal pore size and collective behaviour, suggesting the overall pattern in the complete data set was driven by the data from the two parent species.

4.5.3 Artificial sensor simulations

Analysis of simulated flow behind a 100 mm diameter cylinder (used to generate a wake similar to one generated by a swimming fish (Liao 2007; Fish 2010) at a flow speed of 0.5 m/s (Figure S4.6) revealed areas with substantially higher levels of variation in flow speed (Figure 4.5). Our computational model predicted that the location with the greatest flow speed variation was 200 mm behind the cylinder, and 40 mm offset from its centreline (Figure 4.5A). Data were extracted from this point, initially with no canal structure, showing the flow speeds that a superficial neuromast hair cell would experience. This data gave an oscillating waveform of set period (Figure 4.5B) that is a result of the repeated changes in flow velocity that occur when flow is slowed or even reversed on one side of each successive shed vortex. The waveform oscillated around a mean flow speed of 0.4 m/s, slightly slower than the background flow speed set in the simulation, which is due to the flow being slower in the vicinity behind the cylinder. Flow speed data taken from the same location, but from within the simulated canal-like sensor showed a large decrease in the mean flow speed (Figure 4.5B). This decrease shows the filtering effect that the sensor has on background flow speed, allowing it to measure only the velocity that is induced as a result of the pressure differential between the pores. The mean value of the flow velocity in the artificial canal sensor is close to 0 (Figure 4.5B). Such low velocity may be below the threshold needed to illicit a response from the visual tracker in a physical set-up. However, through additional preliminary trials conducted with a physical 3D printed sensor, we demonstrated that the vortex created by a passing cylinder was able to be detected by a sensor with 10 mm pores and larger (Figure S4.7).

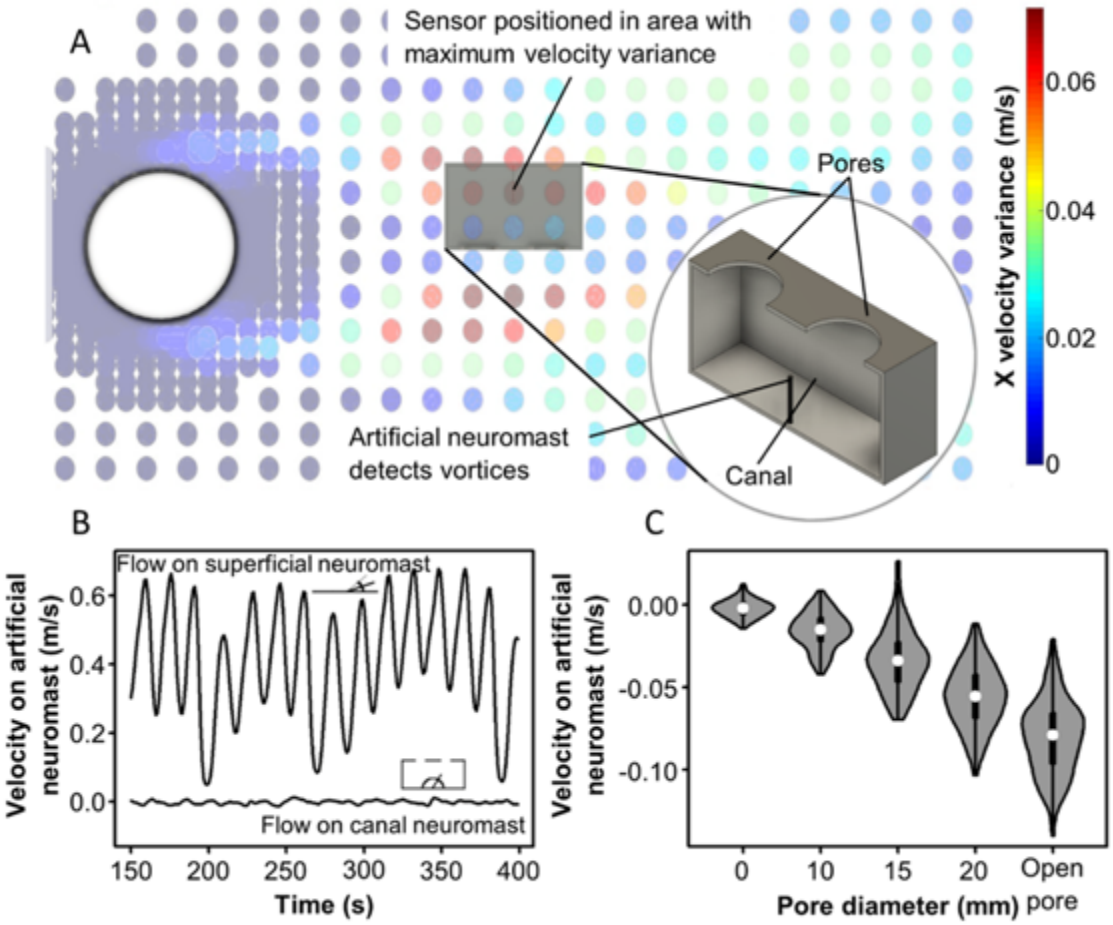


Figure 4.5 The design of the simulated artificial lateral line sensor and its location and response in the simulated environment. (A) Flow velocity variation within the simulated flow tank during a 400 second simulation, showing areas of high levels of flow velocity variation (red) and low levels of flow velocity variation (blue). Each coloured dot in simulation marks a mesh point. The sensor was positioned at the point with highest variation. The position and CAD (computer-aided design) model for the artificial sensor used in simulations are shown. The CAD model is inspired by the structure of a section of canal lateral line, with a single neuromast being positioned between two pores. (B) Comparison of flow velocities seen in flow where no structure is present (upper line) and flow velocities within the canal structure with pore size 15mm (lower line) (shown in A). Schematic drawings of the predicted response that a neuromast of similar stiffness would show without the canal structure (like a superficial neuromast) (top left) and with it (bottom right) are included. (C) Distribution of the speeds seen within the artificial canal with varying sizes of pore, where wider regions of each violin represent more time with flow at the associated velocity, with the black bar representing the interquartile range and the white dot representing the mean. 'Open pore' indicates that there are no pores and instead the entire surface bearing pores (seen in A) is removed to give a single maximum-size pore. Velocities are negative as the sensor is positioned at the point where flow is reversed.

Our simulations of how pore size impacted the sensitivity of the sensor showed that as the diameter of the pore increased, so did the flow velocity acting on the artificial neuromast (Figure 4.5C). There was also a trend of increasing variance in velocities experienced by the neuromast, although these velocities were still well below what a superficial neuromast could be expected to

experience (Figure 4.5C). Given that the responses for all of the different pore sizes are to the same stimuli (i.e. the wake behind a 100 mm cylinder), we can conclude that larger pore sizes result in a greater response. As the strength of the stimulus decreases, the response from each pore size in turn will lessen accordingly (responses from 5 mm pores will become undetectable first, and then 10 mm etc.) As such, larger pore sizes in an artificial lateral line sensor are better able to detect upstream stimuli such as the vortices from an upstream fish. The results here were echoed in the preliminary trials with the 3D printed physical sensor, where 5 mm pores showed no response to the passing vortex, but 20 mm pores showed a strong response (Figure S4.7).

4.6 Discussion

4.6.1 Segregation of lateral line morphology in hybrid fish

Together, the hybrids exhibited a broad range of morphological and behavioural diversity between the two parental phenotypes of 'wide' canals and 'narrow' canals, with most traits, on average, being intermediate between the two (Figure 4.3). These results provide additional evidence of the heritability of several lateral line components, such as pore size and neuromast number (Wark et al. 2012; Chapter 3). Similarly, the collective behaviour of hybrids, as measured by our various metrics of group cohesion, was intermediate between the two parental phenotypes, contributing additional evidence to the existing literature that shoaling in fish has a heritable component (Huizinga et al. 2009; Greenwood et al. 2013; Tang et al. 2020).

Using hybrids from an F_2 interspecific cross enabled us to partially overcome difficulties of testing two distinct species where there is the risk that differences in behaviour arise from distinct species-specific behaviours, or information from other sensory systems, such as olfactory or visual abilities. Given the known roles of these senses, particularly vision, in shoaling and schooling behaviours, it is important that efforts are made to control them such that any variation in behaviour can be better attributed to the lateral line. Our two parental fish live in similar conditions, close to the substrate at a depth of around 10 m. The primary difference in habitat is that *A. stuartgranti* lives over sandy substrate while the *O. lithobates* lives over rocky substrate (Konings 2016). As such we may assume that they possess broadly similar visual systems. While we acknowledge the important role of vision in shoaling, we assume that all hybrids possess similar levels of visual acuity, and we ensure consistent visual conditions across all trials, so we can attribute any changes in observed behaviour to variation in their lateral line system morphologies.

Although vision is likely to be the main sensory modality used to mediate collective behaviour in many fish species (Partridge and Pitcher, 1980; Kowalko et al. 2013; Strandburg-Peshkin et al. 2013; Pita et al. 2015; Chamberlain and Ioannou, 2019), our results provide further evidence that information from the lateral line system contributes to collective behaviour. While it has long been known that the lateral line system is important in informing collective behaviour in fish (Pitcher et al. 1976; Partridge and Pitcher, 1980; Faucher et al. 2010), and that superficial and canal neuromasts differ in the sensitivity to different water movement characteristics (Kroese and Schellart, 1992), the specific roles of the different regions of lateral line system were yet to be established.

4.6.2 Associations between lateral line morphology and cichlid collective behaviour

Here, we found that larger anterior canal pores were associated with a decreased average proximity to nearest neighbours (Figure 4.4) and a decreased group radius. Larger pores offer the canal system greater exposure to external flow conditions, while filtering hydrodynamic background noise, and as such provide fish with more information about their surroundings and the potential for better control responses. In addition, flow around anterior structures is less disrupted by the body of the fish, boundary layer effects and self-generated flows, particularly from the pectoral fins (Tang et al. 2020), making incoming information more accurate than other regions of the body, such as the trunk lateral line. This is supported by evidence that the undulatory swimming motions of fish are optimised to reduce self-generated pressures around the head (Akanyeti et al. 2016), and the importance of corollary discharge in the lateral line nerve (Skandalis et al. 2021). Additionally, fish can respond to a stimulus more quickly via cranial structures, as rheotactic behaviour will expose the head to upstream stimuli first. Studies have shown that fish gain many benefits from shoaling in tighter formations (Magurran, 1990; Pavlov and Kasumyan, 2000; Ioannou et al. 2019; Romenskyy et al. 2020) and as such, improved control and acuity can aid with this.

Fish with more visible anterior and/or posterior superficial neuromasts tend to form less cohesive groups. As superficial neuromasts detect flow velocities (Kroese and Schellart, 1992), they could give broad warning of obstructions (like companions) ahead that have reduced flow velocity behind them. More superficial neuromasts may make individuals more sensitive to these reduced velocity areas and naturally distance themselves from them, resulting in more widely spaced fish. Increased numbers of superficial neuromasts also increase susceptibility to noise, resulting in a greater margin of error in estimating distances to shoal mates, and potentially causing fish to stay further apart. The propensity to remain more distant when uncertain could be to mitigate possible

collisions. Previous work has highlighted the importance of lateral line mechanoreception in preventing collisions within a shoal, with some lateral line-ablated fish colliding with their shoal companions with sufficient force to stun themselves (Pitcher et al. 1976; Partridge and Pitcher, 1980; Faucher et al. 2010). Collisions can have significant negative effects for the individual so an increased sensitivity to slower flow regions may act to avoid physical contact with shoal mates.

Shoaling and schooling are important behaviours that serve multiple purposes in large numbers of teleost fishes (Shaw 1962; Hemmings 1966; Magurran 1990; Pitcher 1983; Pitcher 1986; Tien et al. 2004; Fish 2010; van Oosterom et al. 2016; Ioannou 2017; Ginnaw et al. 2020). To govern the complex dynamics seen in these systems (Tien et al. 2004; Ioannou 2017), evidence suggests that vision (Shaw 1962; Pitcher 1983; Pitcher 1986; Ginnaw et al. 2020), the lateral line system (Pitcher et al. 1976; Partridge and Pitcher, 1980; Faucher et al. 2010), audition (van Oosterom et al. 2016) and olfaction (Hemmings 1966) can be involved, making it difficult to understand the distinct roles they play. Additionally, recent work shows that genetics can play a role, with mutations in certain genes affecting shoaling behaviour in Zebrafish (Tang et al. 2020). On the other hand, the lateral line system has been shown to be essential in regulating not just shoaling (Pitcher et al. 1976; Partridge and Pitcher, 1980; Faucher et al. 2010), but also rheotaxis (Montgomery et al. 1997; Baker and Montgomery, 1999; Olszewski et al. 2012; Suli et al. 2012; Simmons et al. 2013), aggressive interactions with conspecifics (Butler and Maruska, 2016; TelMarsch and Ward, 2020), mating (Medina et al. 2013; TelMarsch and Ward, 2020) and prey detection (Hoekstra and Janssen, 1986; Janssen et al. 1995; Rowe et al. 2003; Schwalbe et al. 2012; Schwalbe and Webb, 2014; Schwalbe et al. 2016). These behaviours may drive lateral line adaptations to differing degrees and the relative importance of each will depend on the ecology of the individual species and populations. For example, cave-living eyeless forms of the Mexican tetra *Astyanax mexicanus* have almost entirely lost their shoaling behaviours and, despite increased numbers of neuromasts, lateral line ablation does not significantly affect this (Kowalko et al. 2013), suggesting their lateral line system has adapted for other functions. In firehead tetras, however, ablation of the lateral line renders an individual unable to shoal normally (Faucher et al. 2010). It has also been shown that the mandibular canal is correlated with feeding behaviour and diet in cichlids (Schwalbe et al. 2012; Schwalbe and Webb, 2014; Schwalbe et al. 2016; Edgley and Genner, 2019). As such, adaptations of the lateral line may not primarily be driven by shoaling, but as we show, variation in the lateral line morphology of the individuals we tested does interact with their collective behaviour.

We hypothesised that superficial neuromasts would be the primary correlates of collective behaviour, with anterior canal pores taking a smaller role. Our results partially support both predictions, with superficial neuromasts mediating group radius, and increased numbers resulting in less cohesion. This is in line with research showing that removal of superficial neuromasts of the trunk results in the target individual becoming more cohesive with the group (Pitcher et al. 1976; Partridge and Pitcher, 1980). We found that the role of the anterior canal neuromasts was greater than expected, with larger canal pores in this region resulting in decreased nearest neighbour distances. It is possible therefore that superficial neuromasts regulate repulsion from neighbours while canal neuromasts regulate attraction.

4.6.3 Designing and testing a novel artificial lateral line sensor

Our modelling of an artificial lateral line demonstrated that a single artificial neuromast within a section of canal was better able to filter out background flow than an artificial superficial neuromast (Figure 4.5), highlighting the important functional differences between the two sub-modalities (Kroese and Schellart, 1992). Additionally, the reduced flow speeds within the sensor allow for the use of a more flexible neuromast, resulting in greater potential deflection in response to low-velocity flows within the sensor, thereby increasing sensitivity. In contrast, a more flexible neuromast in the open flow will deflect significantly from the speed of the background flow, with the potential that maximum deflection will occur, lessening or preventing responses to other stimuli, such as from an upstream fish, from being recorded. Previous work has shown that flow of around 10cm/s is enough to mask the vibrations caused by a vibrating sphere in both still-water (goldfish) and riverine (rainbow trout) fish (Engelmann et al. 2002). Increasing the stiffness of the superficial neuromast might prevent maximum deflection at the same flow velocity, in turn preventing saturation, but by doing so, less deflection will be recorded in response to flow stimuli, reducing the sensitivity of the artificial lateral line.

We also found that a larger pore size resulted in both an increased mean flow velocity and an increased range of flow velocities. The latter entails greater deflections in response to vortices. It also implies that the threshold velocity for the sensor to be able to detect vortices is lower. However, the increasing mean flow velocity showed there was less filtering of the background flow at larger pore size. These trends have been demonstrated before with artificial canal neuromast sensors that were kept at the expected biological scales, but we show that these trends are also present in a macro scale sensor. Further work has been undertaken to fully optimise the sensor for use in swarm robotics applications (Scott 2022). It should be noted that the low internal

flow velocity does have the potential to reduce the efficacy of the sensor, but that the future work addresses and overcomes this (Scott 2022). At this stage, the design is only intended to be a proof of concept that confirms the observations and shows that a macro-scale canal can be effective.

It should be noted that the bony pores of cichlids are covered with a thin membrane that limits flow entering the canals; only a small hole exists that fluid can pass through (Becker et al. 2016). A larger area of membrane however can act to amplify external signals making them more perceptible within the canals (Becker et al. 2016). Our sensor design adopted this principle but attempted to simplify it by removing the membrane, which enabled us to explore if it was still effective at detecting pressure changes. Larger pores here resulted in an increase in both the mean flow velocity and the range of velocities detected, indicating that with increasing pore size, the sensor gives a greater response to an external stimuli. The external stimuli are the same in each case, i.e. the vortex street.

Significant work has already been done in the field of artificial lateral lines (reviewed by Hu et al. 2019), but the majority of these systems use Micro-Electrical-Mechanical-Systems (MEMS) (Fan et al. 2002; Zhou & Liu, 2008; Kottapalli et al. 2014), remaining true to the scales of the biological systems they are based on. This is because flow properties change significantly as scales do. However, MEMS can be difficult to mass produce (Shoaib et al. 2016). Some artificial lateral lines that use off-the-shelf pressure or optical sensors exist (Ji et al. 2017; Zheng et al. 2018) but these use multiple sensors per lateral line to effectively localise a source, increasing cost and complexity. In this work, we show that a single instance of a macro-scale sensor can effectively filter background flow velocity and output a sinusoidal signal (similar to the input). With further work, this could be used to help underwater swimming robots to locate and follow other swimming robots without the need for visual input. This is especially helpful in the underwater environment which is often dark or turbid. We also show that larger scale artificial canal lateral line sensors are feasible, something that has not been explored, which could open new areas of research in artificial lateral line design. Additionally, the simplicity of the proposed design makes it easy and inexpensive to mass produce (Figure S4.7). As such, it will be useful in the field of swarm robotics, where the numbers involved prohibit anything expensive or difficult to manufacture. Swarms of underwater swimming robots could then be used for environmental monitoring or search and rescue, where the large numbers improve area coverage and parallel processing. Several systems exist that use an artificial lateral line to gain information about a neighbouring individual (Zheng et al. 2020a; Zheng et al. 2020b; Zheng et al. 2019; Li et al. 2022) which mark the first steps towards an artificial lateral

line-controlled swarm, and this work could help to reduce the cost and complexity of the lateral line required.

4.6.4 Conclusions

Our results have implications for understanding the lateral line system of fishes in the context of their evolution, ecology, and behaviour. We demonstrate how to quantify the morphological disparity within multiple lateral line system structures, and to test the consequences of this variation on collective behaviour. Our results also highlight how different aspects of lateral line system morphology differentially affect shoaling behaviour, and specifically identify the importance of the head structures in shoaling. This could help to explain how globally-important shoaling species like herrings, shads and sardines, that have well developed anterior lateral line canal morphology, albeit different to cichlids, are able to coordinate their collective movement in the absence of posterior lateral line canals (Jørgensen, 1985; Stephens, 1985). In addition, these findings can inform the design of artificial lateral lines. Biomimetic artificial lateral line systems already exist (Fan et al. 2002; Zhou & Liu, 2008; Kottapalli et al. 2014; Hu et al. 2019), but these are currently still complex and can be expensive to mass produce. The design and the simulation results presented here could help to develop a simple, inexpensive, macro-scale sensor capable of informing minimalistic underwater robots for swarming purposes, eventually without the need for visual signals.

4.7 Supplementary Information

Table S4.1. Mean and standard deviation of phenotypes of the focal individuals included in this study: Head length and standard length; behavioural response variables; neuromast counts; and mean pore area measurements.

	All		**Hybrids**		***A. stuartgranti***		***O. lithobates***	
Sample size	116		86		10		10	
	μ	σ	μ	σ	μ	σ	μ	σ
Standard length (mm)	116.71	16.97	117.31	17.71	110.50	10.28	117.10	13.41
Head length (mm)	29.21	3.91	29.43	4.07	27.09	2.27	29.23	2.92
Nearest neighbour (mm)	93.98	61.27	90.77	50.32	66.14	19.40	158.78	123.42
Coefficient of variation (mm)	0.95	0.64	0.94	0.62	0.67	0.22	1.42	0.94
Nearest neighbour upstream (mm)	123.65	63.96	121.92	53.39	86.92	27.89	182.77	125.84
Coefficient of variation upstream (mm)	0.96	0.45	0.98	0.42	0.63	0.16	1.16	0.66
Nearest neighbour field of view (mm)	119.90	75.49	118.57	64.85	73.59	27.47	185.38	142.35
Coefficient of variation field of view (mm)	1.08	0.55	1.10	0.52	0.68	0.22	1.30	0.81
Group radius (mm)	211.75	77.20	214.37	70.05	117.94	60.94	221.59	135.79
Head superficial neuromasts	73	24.8	72	23.5	58	25.5	99	16.7
Trunk superficial neuromasts	146	37.5	150	37.3	131	34.0	114	19.4
Trunk canal neuromasts	28	2.4	28	2.5	26	1.8	27	1.7
Upper trunk canal neuromasts	59	15.2	61	15.4	51	12.1	56	12.9
Lower trunk canal neuromasts	17	1.9	17	1.9	16	1.7	18	1.4
Caudal fin superficial neuromasts	82	50.3	83	53.1	71	30.8	78	34.5
Mean mandibular canal pore area (mm^2)	0.258	0.093	0.251	0.069	0.432	0.101	0.152	0.029
Mean preopercular canal pore area (mm^2)	0.770	0.201	0.750	0.159	1.149	0.193	0.584	0.089
Mean infraorbital canal pore area (mm^2)	0.368	0.087	0.374	0.062	0.540	0.052	0.236	0.030
Mean otic canal pore area (mm^2)	0.910	0.247	0.875	0.152	1.501	0.232	0.652	0.079

Table S4.2 Model comparisons using the Akaike Information Criterion scores for neuromast models that explain variability in mean nearest neighbour distance, for a dataset that uses all individuals (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Anterior.Canal.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	0.0	-ve	0.0448
scale(Anterior.Canal.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	0.4	-ve	0.0331
scale(Anterior.Superficial.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	5.3	+ve	0.0116
Null	6	5.8	N/A	0
scale(Anterior.Superficial.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	6.1	+ve	0.0176
scale(Posterior.Superficial.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	7.3	+ve	0.0031
scale(Lower.Posterior.Canal.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	7.4	+ve	0.0027
scale(Lower.Posterior.Superficial.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	7.5	+ve	0.0021
scale(Posterior.Canal.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	7.9	-ve	0.0004
scale(Tail.Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	7.9	-ve	0.0002
scale(Lower.Posterior.Canal.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	9.4	-ve	0.0034
scale(Posterior.Superficial.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	9.4	-ve	0.0031
scale(Lower.Posterior.Superficial.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	9.4	-ve	0.0031
scale(Posterior.Canal.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	9.8	-ve	0.0016
Scale(Tail.Neuromasts)*Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	8	10.0	+ve	0.0004

Table S4.3 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean nearest neighbour distance**, for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	0.0	-ve	0.0255
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.2	+ve	0.0190
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.4	-ve	0.0296
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.0	-ve	0.0265
Null	6	2.7	N/A	0
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial.Order+(1|ID)	7	4.4	+ve	0.0023
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.6	+ve	0.0012
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.8	-ve	0.0003
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.8	+ve	0.0001
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.8	+ve	1e-05
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.0	-ve	0.0053
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.7	-ve	0.0015
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.8	+ve	0.0013
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.8	-ve	0.0011
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.9	-ve	0.0005

Table S4.4 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean group radius**, for the dataset that uses **all individuals** (hybrids and parent species), where **total body length is included as a covariate term** in all models. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	+ve	0.0299
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.9	+ve	0.0302
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	3.6	+ve	0.0145
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.1	-ve	0.0124
Null	6	5.0	NA	
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	5.5	+ve	0.0152
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.1	-ve	0.0124
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	6.6	-ve	0.0017
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	6.8	-ve	0.0007
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	6.9	+ve	0.0004
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	6.9	-ve	0.0004
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	8.4	-ve	0.0023
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	8.6	+ve	0.0016
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	8.7	-ve	0.0013
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	8.9	+ve	0.0005

Table S4.5 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean group radius**, for a dataset that uses **only hybrids**. Total length is included as a covariate term. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	+ve	0.0249
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.9	+ve	0.0154
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.9	-ve	0.0153
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.0	+ve	0.0249
Null	6	2.8	NA	0
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.8	+ve	0.0157
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.9	-ve	0.0154
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.1	-ve	0.0036
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.4	-ve	0.0020
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.6	-ve	0.0012
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.8	+ve	0.0001
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.1	-ve	0.0036
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.4	-ve	0.0020
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.6	-ve	0.0012
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	6.8	-ve	0.0002

Table S4.6 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean nearest upstream neighbour distance**, for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	0.0	-ve	0.0698
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	5.8	-ve	0.0364
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	7.6	+ve	0.0289
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	8.0	+ve	0.0364
Null	6	12.1	N/A	0
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	12.2	+ve	0.0091
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	12.9	+ve	0.0059
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	13.5	+ve	0.0030
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	14.1	+ve	0.0099
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	14.2	-ve	0.0002
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	14.2	-ve	0
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	15.0	+ve	0.0061
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	15.7	+ve	0.0031
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	15.9	-ve	0.0019
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	16.2	+ve	0.0006

Table S4.7 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean nearest upstream neighbour distance**, for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Associations	LRT
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	0.0	-ve	0.0595
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.0	+ve	0.0278
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	4.5	-ve	0.0254
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	5.9	-ve	0.0291
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	6.5	+ve	0.0150
Null	6	7.1	N/A	0
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	8.4	+ve	0.0045
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	8.6	+ve	0.0150
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	9.3	-ve	1e-05
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	9.3	+ve	0
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	9.3	-ve	0
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	10.4	+ve	0.0059
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	11.1	+ve	0.0018
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	11.3	-ve	0.0008
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	11.5	+ve	1e-05

Table S4.8 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean nearest neighbour distance within 'field of flow detection',** for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	8	-ve	0.0503
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.4	7	-ve	0.0354
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.7	7	+ve	0.0296
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.8	8	+ve	0.0385
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.5	7	+ve	0.0131
Null	7.4	6	N/A	0
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.6	7	+ve	0.0084
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8.4	7	+ve	0.0049
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8.6	8	+ve	0.0134
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	9.4	7	+ve	0.0004
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	9.5	7	+ve	0
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	9.8	8	+ve	0.0084
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	10.3	8	-ve	0.0062
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	11.2	8	-ve	0.0020
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	11.3	8	+ve	0.0016

Table S4.9 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean nearest neighbour distance within 'field of flow detection',** for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	7	+ve	0.0292
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.5	8	-ve	0.0328
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.7	7	-ve	0.0206
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.0	8	-ve	0.0301
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	3.2	7	+ve	0.0126
Null	3.4	6	N/A	0
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.4	7	+ve	0.0062
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.1	8	+ve	0.0139
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.4	7	+ve	0.0007
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.4	7	+ve	0.0005
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.5	7	-ve	0
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.3	8	+ve	0.0078
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.5	8	+ve	0.0011
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.6	8	-ve	0.0009
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.6	8	+ve	0.0007

Table S4.10 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean distance to centroid,** for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	8	-ve	0.0441
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.8	7	-ve	0.0318
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.4	7	+ve	0.0247
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.3	8	+ve	0.0259
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.7	7	+ve	0.0106
Null	5.9	6	N/A	0
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.7	7	+ve	0.0058
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.1	7	+ve	0.0042
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.7	8	+ve	0.0109
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8.0	7	-ve	0.0004
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8.0	7	+ve	0.0001
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8.9	8	+ve	0.0059
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	9.3	8	+ve	0.0042
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	9.8	8	-ve	0.0016
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	10.1	8	+ve	0.0005

Table S4.11 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean distance to centroid,** for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	8	-ve	0.0264
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.6	7	+ve	0.0235
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.7	8	+ve	0.0339
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.1	8	-ve	0.0323
Null	2.8	6	N/A	0
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.9	7	+ve	0.0113
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.1	7	+ve	0.0047
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.4	8	+ve	0.0151
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.8	7	+ve	0.0009
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.0	7	-ve	0
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.0	7	-ve	0
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.2	8	+ve	0.0054
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.0	8	+ve	0.0010
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.0	8	+ve	0.0009
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.2	8	-ve	0

Table S4.12 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean bounding hull size,** for a dataset that uses **all individuals**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	7	+ve	0.0252
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.2	7	+ve	0.0242
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.1	8	+ve	0.0298
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.9	8	+ve	0.0264
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.9	7	-ve	0.0128
Null	3.6	6	N/A	0
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.8	8	-ve	0.0135
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.0	7	+ve	0.0032
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.7	7	-ve	0.0003
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.8	7	+ve	0
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	5.8	7	+ve	0
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.3	8	+ve	0.0072
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.7	8	+ve	0.0054
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.6	8	+ve	0.0014
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.8	8	-ve	0.0003

Table S4.13 Model comparisons using the Akaike Information Criterion scores for **neuromast** models that explain variability in **mean bounding hull size,** for a dataset that uses **only hybrids.** Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Anterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	7	-ve	0.0261
scale(Anterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.4	7	+ve	0.0242
scale(Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.1	7	+ve	0.0206
scale(Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.1	8	+ve	0.0312
scale(Anterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.1	8	-ve	0.0265
scale(Anterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.4	8	+ve	0.0251
Null	2.8	6	N/A	0
scale(Lower_Posterior_Superficial_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.6	7	+ve	0.0020
scale(Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.6	7	-ve	0.0016
scale(Lower_Posterior_Canal_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.9	7	+ve	0
scale(Tail_Neuromasts)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.9	7	-ve	0
scale(Tail_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.2	8	+ve	0.0049
scale(Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.4	8	-ve	0.0036
scale(Lower_Posterior_Superficial_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.5	8	+ve	0.0035
scale(Lower_Posterior_Canal_Neuromasts)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7.1	8	+ve	0.0001

Table S4.14 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean nearest neighbour distance**, for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	-ve	0.0269
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.6	-ve	0.0241
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.5	-ve	0.0206
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.7	-ve	0.0288
scale(Dentary_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.7	-ve	0.0194
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.9	-ve	0.0281
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.4	-ve	0.0257
scale(Dentary_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.3	-ve	0.0217
Null	6	4.0	N/A	0

Table S4.15 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean nearest neighbour distance**, for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	-ve	0.0134
Null	6	0.4	N/A	0
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.6	-ve	0.0049
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.8	+ve	0.0155
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.0	-ve	0.0027
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.4	-ve	0.0005
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.7	+ve	0.0056
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.2	-ve	0.0029
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.6	-ve	0.0007

Table S4.16 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean group radius**, for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	-ve	0.0183
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.2	-ve	0.0173
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.7	-ve	0.0201
Null	6	2.0	N/A	0
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.4	-ve	0.0173
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	3.2	-ve	0.0045
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	3.3	-ve	0.0038
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	5.1	-ve	0.0054
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	5.1	-ve	0.0052

Table S4.17 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean group radius**, for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Association	LRT
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	-ve	0.0120
Null	6	0.1	N/A	0
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.5	-ve	0.0094
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.4	-ve	0.0044
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.9	+ve	0.0135
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.2	-ve	0.0004
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.6	+ve	0.0098
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.6	+ve	0.0044
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.3	-ve	0.0005

Table S4.18 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean nearest upstream neighbour distance**, for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Associations	LRT
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	-ve	0.0337
scale(Otic_Canal_Pore)+Turbulent/Laminar+scale(Length) +Trial_Order+(1|ID)	7	0.0	-ve	0.0336
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.9	-ve	0.0253
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.0	-ve	0.0341
Scale(Otic_Canal_Pore)*Turbulent/Laminar+scale(Length) +Trial_Order+(1|ID)	8	2.2	+ve	0.0336
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.4	-ve	0.0235
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.1	-ve	0.0254
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.3	-ve	0.0243
Null	6	5.6	N/A	0

Table S4.19 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean nearest upstream neighbour distance**, for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Associations	LRT
Null	6	0.0	N/A	0
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.1	-ve	0.0110
Scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	0.4	+ve	0.0210
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.2	-ve	0.0050
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	1.7	+ve	0.0140
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.9	+ve	0.0016
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.9	-ve	0.0013
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.9	+ve	0.0026
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.1	-ve	0.0013

Table S4.20 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean nearest neighbour distance within 'field of flow detection'** for a dataset that uses **all individuals** (hybrids and parent species). Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Associations	LRT
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	0.0	-ve	0.0399
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.5	-ve	0.0337
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.1	+ve	0.0402
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.7	-ve	0.0285
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.8	-ve	0.0279
Scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.4	+ve	0.0346
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.3	+ve	0.0308
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	4.9	+ve	0.0283
Null	6	7.1	N/A	0

Table S4.21 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean nearest neighbour distance within 'field of flow detection'** for a dataset that uses **only hybrids**. Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	df	dAICc	Associations	LRT
Null	6	0.0	N/A	0
Scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	0.1	+ve	0.0224
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	0.9	+ve	0.0181
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.1	-ve	0.0056
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.7	-ve	0.0026
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	1.9	+ve	0.0016
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	7	2.0	-ve	0.0009
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	2.9	+ve	0.0075
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	8	3.3	+ve	0.0055

Table S4.22 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean distance to centroid,** for a dataset that uses **all individuals.** Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	7	-ve	0.0248
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.1	7	-ve	0.0200
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.1	8	-ve	0.0250
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.6	7	-ve	0.0136
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.7	8	-ve	0.0223
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	3.2	7	-ve	0.0110
Null	3.5	6	N/A	0
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.6	8	-ve	0.0141
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.7	8	-ve	0.0139

Table S4.23 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean distance to centroid,** for a dataset that uses **only hybrids.** Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	7	-ve	0.0136
Null	0.4	6	N/A	0
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.2	7	-ve	0.0071
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.4	8	-ve	0.0177
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.0	7	-ve	0.0028
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.4	7	+ve	0.0011
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.8	8	-ve	0.0105
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.1	8	-ve	0.0035
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.5	8	+ve	0.0011

Table S4.24 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean bounding hull,** for a dataset that uses **all individuals.** Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
scale(Otic_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.0	7	-ve	0.0205
scale(Preopercular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.5	7	-ve	0.0138
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.1	8	-ve	0.0205
Null	2.5	6	N/A	0
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	3.6	8	-ve	0.0140
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.4	7	-ve	0.0011
scale(Infraorbital_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	4.5	7	-ve	0.0008
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.5	8	-ve	0.0015
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	6.6	8	-ve	0.0008

Table S4.25 Model comparisons using the Akaike Information Criterion scores for **pore morphology** models that explain variability in **mean bounding hull,** for a dataset that uses **only hybrids.** Total length is included as a covariate term. LRT is likelihood-ratio test and indicates how much better a model is than the null model.

Model Name	dAICc	df	Associations	LRT
Null	0.0	6	N/A	0
scale(Otic_Canal_Pores)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	0.9	7	-ve	0.0068
scale(Infraorbital_Canal_Pores)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.2	7	+ve	0.0052
scale(Preopercular_Pore_Canal)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	1.2	7	-ve	0.0050
scale(Mandibular_Canal_Pore)+Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.2	7	+ve	0
scale(Infraorbital_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.4	8	+ve	0.0101
scale(Preopercular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.8	8	-ve	0.0084
scale(Otic_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	2.8	8	-ve	0.0082
scale(Mandibular_Canal_Pore)*Turbulent/Laminar +scale(Length)+Trial_Order+(1|ID)	3.4	8	+ve	0.0050

Figure S4.1 Behavioural experimental setup. (A) CAD model of the final design of the tank, created prior to construction. (B) Image of the tank used in the experiments, taken from downstream. (C). A still image taken from a video recording of the single hybrid individual (located at the back of the shoal) with the four *Hemitilapia oxyrhynchus* companions during a trial with laminar flow. Image is not representative of the general location of fish over the trial and only meant to give an impression of what the trials looked like.

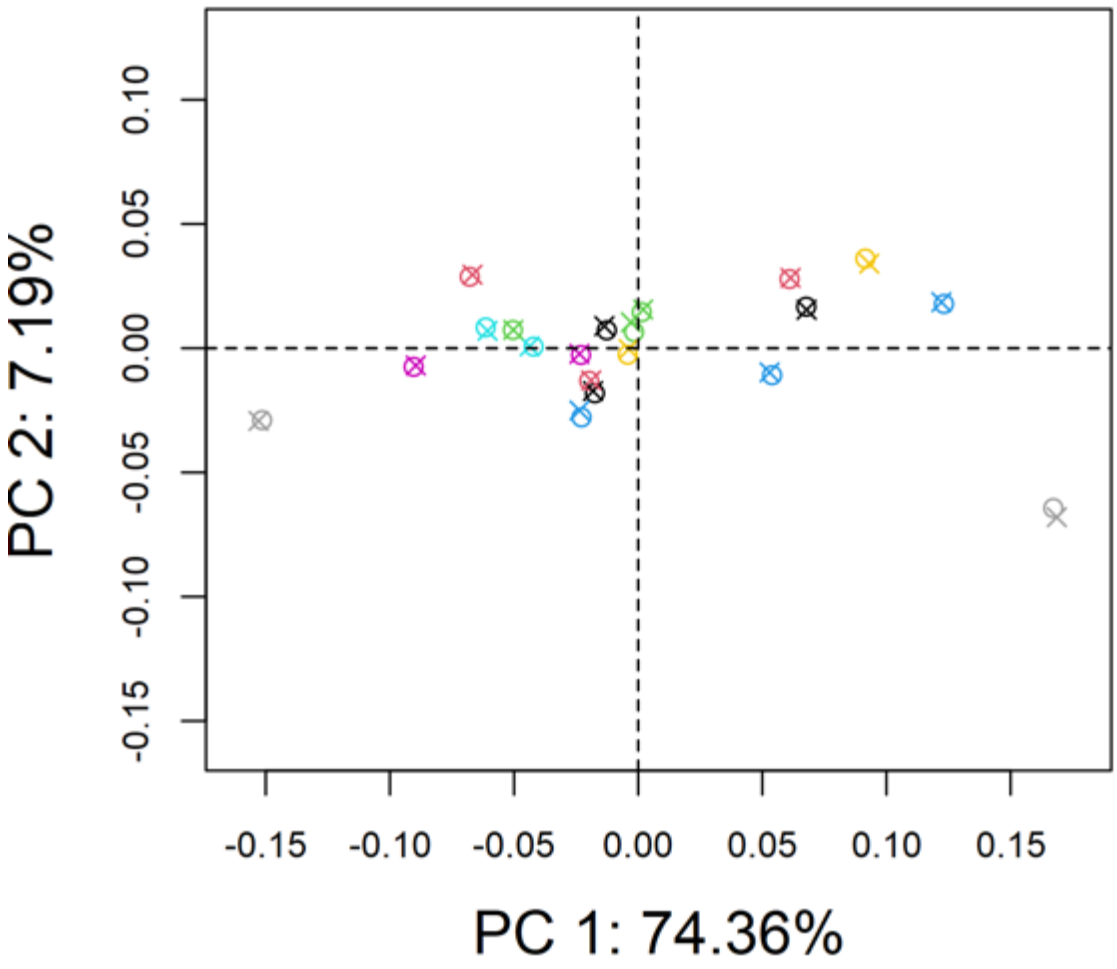


Figure S4.2. Testing for digitisation error in the 2D landmarking process. Prinicipal components analysis on Procrustes coordinates of dentary canal pore landmarks in 20 individuals subjected to repeated 2D landmarking (40 sets of landmarks). Landmarks used are outlined in Figure 4.1. The shape of the point indicates landmarking order (circle = original landmarking; cross = repeated landmarking). Each colour represents a different specimen (Individuals 1-20 in our dataset).

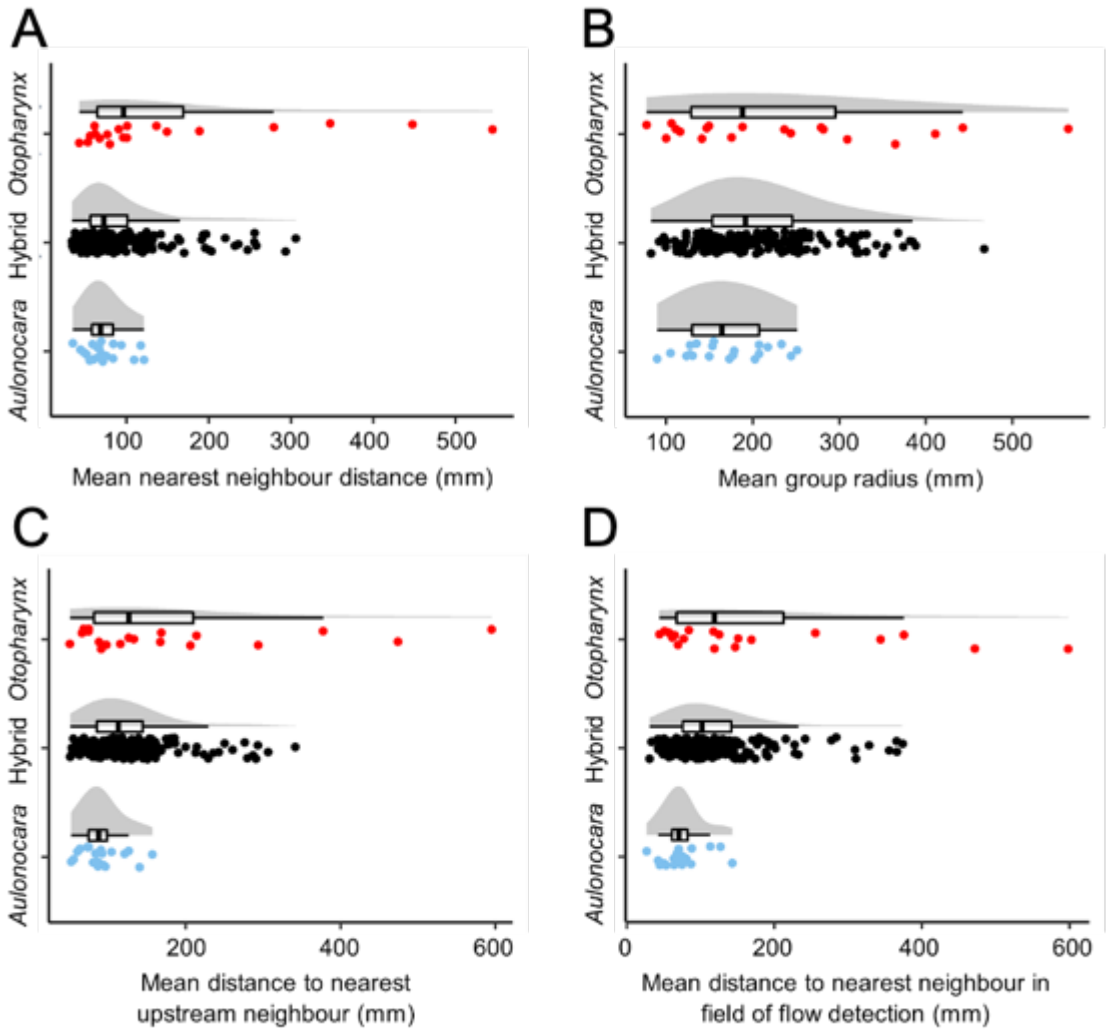


Figure S4.3. Raincloud plots showing the distributions of the number of visible neuromasts in different regions, organised by focal population. (A) Mean nearest neighbour distance. (B) Mean group radius. (C) Mean distance to nearest upstream neighbour. (D) Mean distance to nearest neighbour within the field of flow detection, where the field of flow detection is considered the region both in front of the focal fish and up to 50 mm to either side.

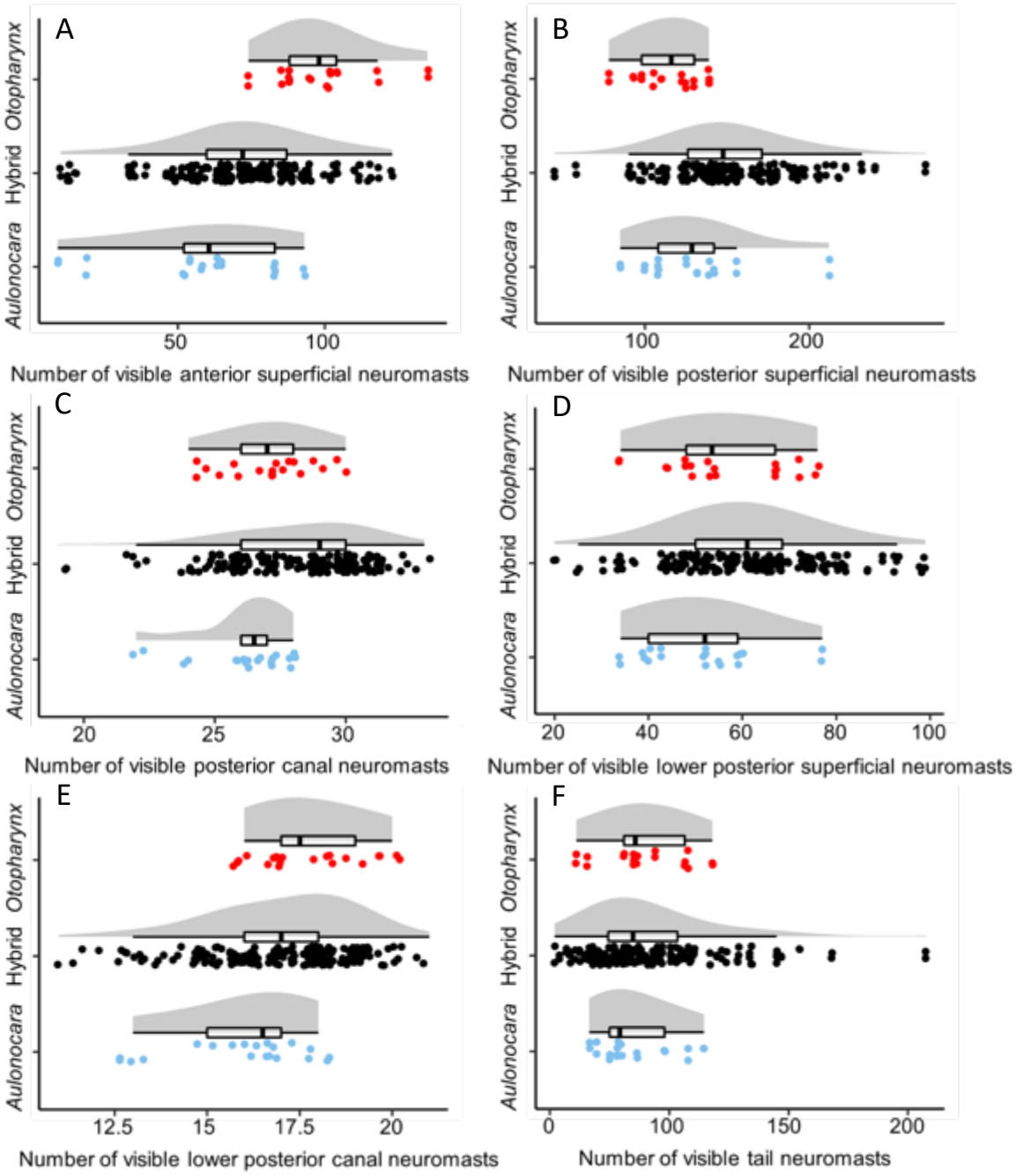


Figure S4.4. Raincloud plots showing the distributions of the number of visible neuromasts in different regions, organised by focal population. (A) Visible superficial neuromasts in the anterior lateral line system. (B) Visible canal neuromasts in the anterior lateral line system. (C) Visible superficial neuromasts in the posterior lateral line system. (D) Visible canal neuromasts in the posterior lateral line system. (E) Visible superficial neuromasts in the lower posterior lateral line system. (F) Visible canal neuromasts in the lower posterior lateral line system. (G) Visible neuromasts in the tail lateral line system.

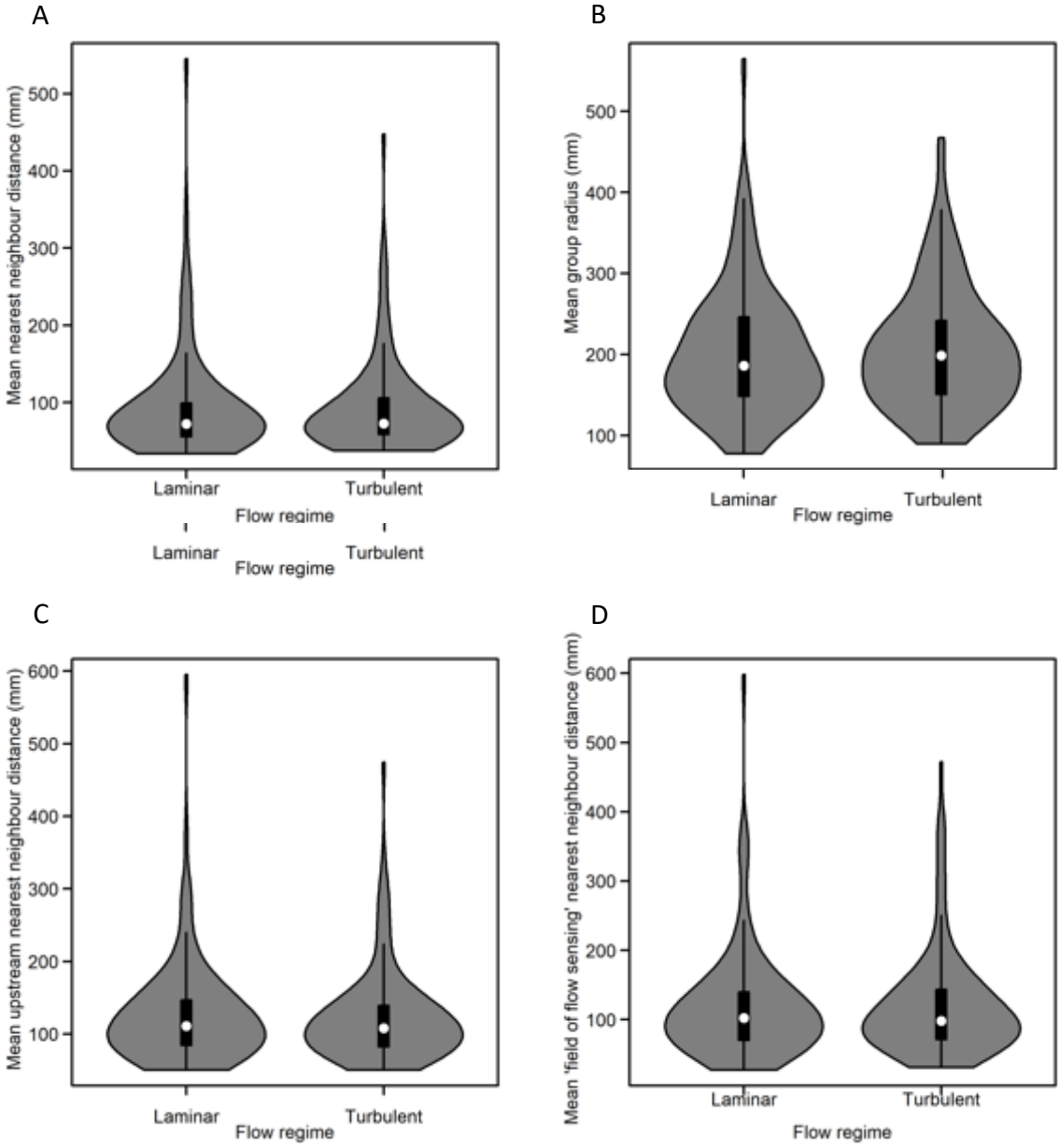


Figure S4.5. Violin plots showing differences in behaviour between laminar and turbulent treatments for (A) nearest neighbour distance, (B) group radius, (C) upstream nearest neighbour distance, and (D) 'field of flow sensing' nearest neighbour distance.

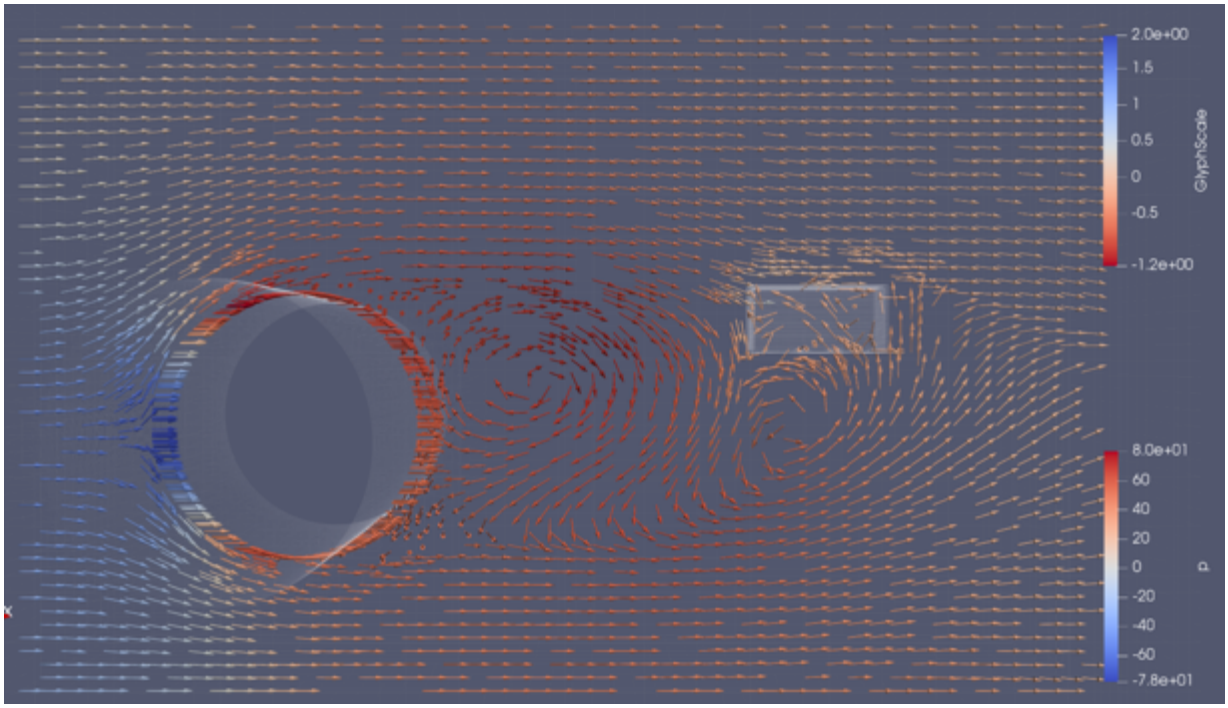


Figure S4.6. Vortices shed from an upstream cylinder are detected by the Artificial Lateral Line (ALL) sensor in simulation. Present in the image are two vortices, rotating in opposite directions, that have been shed alternately from opposite sides of the upstream cylinder. Only a single layer of flow velocities is presented here for clarity.

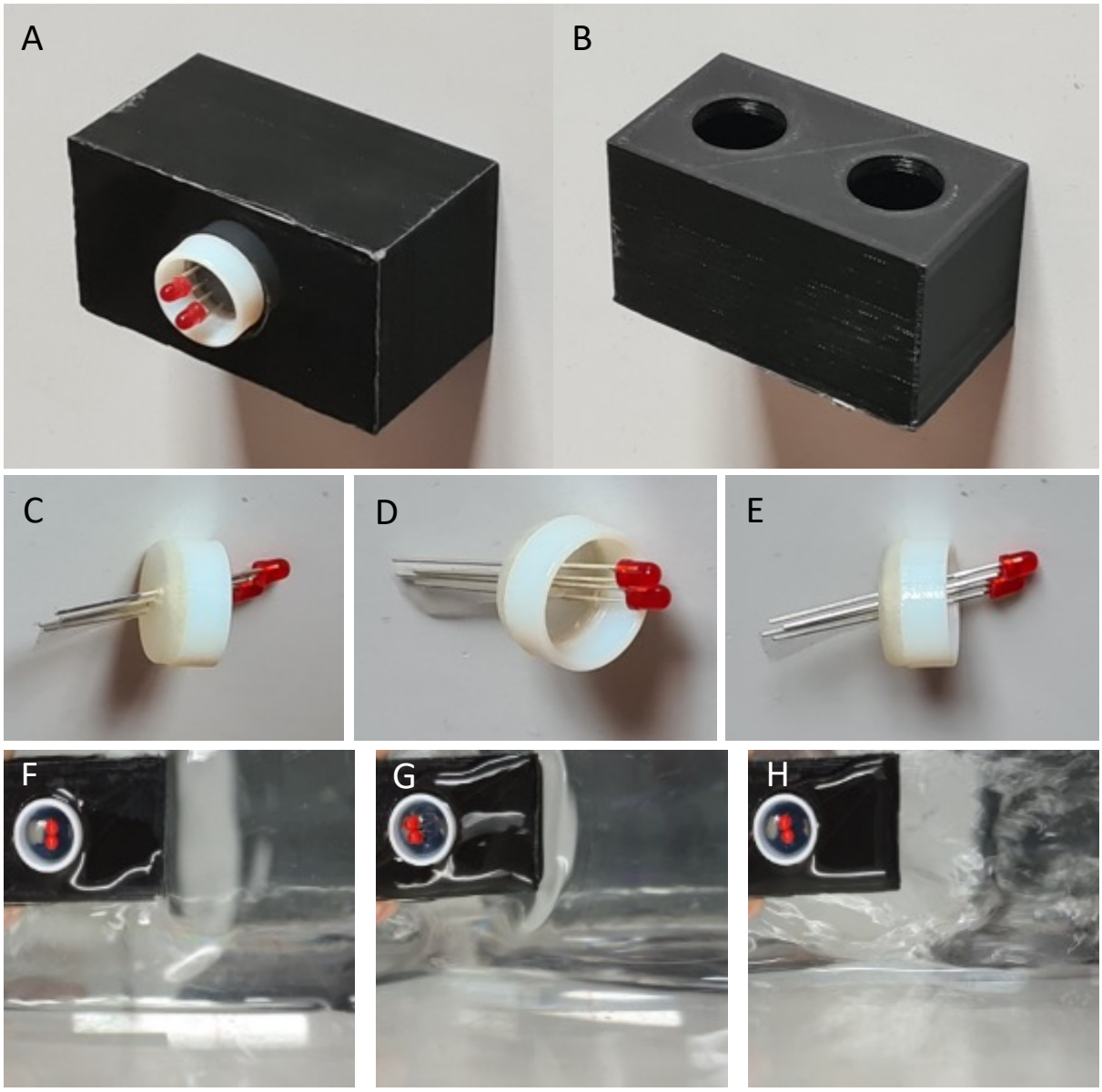


Figure S4.7. A prototype physical sensor based on the CAD designs. (A) View of the back of the sensor, which is the side of the sensor that is monitored by the camera for corresponding deflection. (B) A view of the front of the sensor, which is the side of the sensor that is exposed to the flow being monitored. (C) The removeable module and flexible membrane through which the LED is pierced. The side mainly seen here is fully submerged and moves with the flow. (D) The LED bulbs were filmed and tracked to measure the deflection of the pins. (E) The LEDs were positioned so their midpoint was at the same point as the flexible membrane allowing for equal deflections on either side. (F)-(H) Stills from a video of a cylinder being dragged through the water under the sensor. At the moment of the cylinder's passing (G), the LED recorded a deflection, before returning to neutral (H).

Diving for fish in Lake Masoko, Tanzania, October 2019
Photo: Madeleine Carruthers

Chapter 5

Population-specific habitat-linked directional selection drives cichlid divergent evolution in an East African crater lake

5.1 Abstract

Ecological divergence commonly accompanies speciation events, supporting a role for divergent natural selection during the speciation process. Precisely how selection operates during adaptive radiation, however, is often unclear. Here we describe a study of cichlid fishes in an isolated crater lake in southern Tanzania, Lake Masoko. Based on whole genome sequences and ecomorphological data from individuals collected from known depth bands, we reinforce existing evidence for two genomically and morphologically divergent ecomorphs, occupying shallow littoral and deep benthic habitats respectively. By quantifying the relative fitness of these individuals we find clear bimodal fitness landscapes in both genomic and morphological space, associating with the habitats occupied by each ecomorph. Mapping relative fitness of fish to individual phenotypic traits, we show complex patterns of directional selection on traits within habitat-linked subpopulations that collectively result in a lake-wide pattern of divergent phenotypic evolution between these ecomorphs. We suggest that these diverging patterns of trait fitness in habitat-linked subpopulations have underpinned ecological speciation, and this process is likely to have contributed to the vast cichlid fish radiations within the East African Great Lakes.

5.2 Author Contributions

Duncan Edgley conducted: all data cleaning, compilation and verification; analysis of scale growth data; CT scanning, reconstruction and digitising of specimens; all selection gradient analyses, and writing of the manuscript. This chapter was conceived, analysed and written under the supervision of **Martin Genner**.

Fish were caught over 10 years of sampling by **Martin Genner**, **George Turner**, **Emília Santos**, **Alan Hudson**, **Gregoire Vernaz**, **Richard Durbin**, **Benjamin Ngatunga**, **Semvua Mzighani** and **Asilatu Shechonge**. Light measurements were taken by **Alan Hudson**. Linear morphological measurements, including the lower pharyngeal jaw, but excluding lateral line measurements were taken by **Alexandra Tyers** and **Denise Crampton**. Data cleaning, compilation, and verification of these was conducted by **Duncan Edgley**. Eggspot count and contrast data were collected by **Emília Santos**. Scale circuli were counted by **Giulia Trauzzi**, **Clara Montgomery**, **Jessie Fernando** and **Urte Balseviciute**, all subsequent scales analyses were by **Duncan Edgley**. Fish reared for growth proxy validation and evaluation of the association between body size and gonad weight were reared by **Domino Joyce** and **Alan Smith**. Scale circuli for growth proxy validation were counted by **Dylan Oliver**, & analysed by **Duncan Edgley**. Gonad weight data were collected by **Andrew Saxon** and analysed by **Duncan Edgley**. Whole genome shotgun sequencing, SNP calling and genomic PCA was conducted with **Tyler Linderoth**, **Hannah Munby**, **Richard Durbin** and **Bettina Fischer**. All further analyses were conducted by **Duncan Edgley** and **Martin Genner**, including calculation of selection gradients and GWAS. **Duncan Edgley** and **Martin Genner** wrote the manuscript.

5.3 Introduction

Natural selection is a primary driver of adaptive evolution (Darwin 1859; Coyne and Orr, 2004; Orr 2005). Models of ecological speciation theorise that local environmental conditions drive disruptive selection on ecologically-functional traits, causing differences between populations to accumulate even in the presence of gene flow, ultimately leading to reproductive isolation (Smith 1966; Schluter 2000; Rundle and Nosil, 2005). Ecologically-driven speciation is thought to be a major contributor to species diversity, particularly within adaptive radiations (Schluter 2009; Pfaender et al. 2016), and evidence for speciation between proximate populations with gene flow is now well established (Barluenga et al. 2006; Savolainen et al. 2006; Bolnick and Fitzpatrick, 2007; Malinsky et al. 2015). However, studies linking genetic and morphological evidence of ecological speciation with quantitative evidence for disruptive selection are relatively scarce. Studying associations between traits and fitness at the early stages of speciation should therefore provide a clearer understanding of how selection operates at the microevolutionary level, and how it contributes to the adaptive radiation process.

In the context of ecological speciation, studies often measure selection across partially or fully reproductively isolated species pairs, and tend to focus on key functional traits, either through translocation (Calsbeek and Smith, 2008; Maan et al. 2017), mark-recapture (Calsbeek and Smith, 2008; Hendry et al. 2009) or common-garden experiments (Rajkov et al. 2018). Such studies have found strong correlations between fitness and phenotypes, with individuals typically having higher fitness within their native environments. Similar results have been shown within the broader context of adaptive radiations, where empirical evidence for disruptive selection driving adaptive diversification has been identified in many well-studied systems, such as Galápagos finches (Hendry et al. 2009; Sulloway and Kleindorfer, 2013; Beausoleil et al. 2019), Lake Victoria cichlids (van Rijssel et al. 2018), and San Salvador *Cyprinodon* pupfishes (Martin and Wainwright, 2013; Patton et al. 2022; Martin and Gould, 2020).

Selection-driven divergence in natural environments is often quantified using quadratic selection gradients (Lande and Arnold, 1983; Kingsolver et al., 2001; Brodie et al. 2005) (Figure S5.1). However, during ecological speciation events there is often intraspecific population structure across habitats, and this stratification can associate with distinct genetically differentiated ecotypes. It is possible that analyses aimed at quantifying disruptive natural selection may be hindered by this genetic and/or habitat-linked stratification within populations. An alternative approach,

therefore, would be to measure selection independently within known subpopulations or microhabitats, and compare the patterns of selection between them.

Here we explore contrasting signatures of natural selection among structured populations of the cichlid species *Astatotilapia calliptera* in a small crater lake in Tanzania: Lake Masoko. In this lake, two ecologically distinct populations, or ecomorphs (shallow-water 'littoral' and deep-water 'benthic'), have emerged within the past 1,000 years, despite ongoing gene flow (Malinsky et al. 2015; Munby et al. 2022). These genetically differentiated subpopulations (Figure 5.1H) are distinguished by their male breeding colours (Figure 5.1G) but have also diverged in diet (Figure 5.1K), craniofacial morphology (Figure 5.1I), and in key functional genes (Malinsky et al. 2015; Carruthers et al. 2022; Vernaz et al., 2022). Here, we collect genomic, morphological, and relative fitness data from individuals from known depth bands, to visualise the fitness landscapes for both morphology and genome. In addition, we map phenotype to fitness in morphological traits to gain insight into how selection operates on phenotype during divergence. We use within-ecomorph genome-wide association studies (GWAS) to explore patterns of associations between our fitness proxy and variants in the genome. Collectively the analyses provide insight into how fitness landscapes are generated during the early stages of speciation.

5.4 Materials and Methods

5.4.1 Sample collection

491 *Astatotilapia calliptera* (442 males and 49 females) were caught at Lake Masoko, Tanzania, in three sampling expeditions between 2014 and 2018. Samples were caught using SCUBA and block nets from five depth bands: 0 – 5 m (n=153); 5 – 10 m (n=45); 10 – 15 m (n=116); 15 – 20 m (n=45); and > 20 m (n=132) (Figure 5.1E-G). Depth was confirmed using SCUBA dive gauges. Fish were euthanised with an overdose of MS-222 (Sigma Aldrich), before being pinned and dried. The left side of each fish was photographed for morphometric analyses. Pectoral fin clips were taken for whole-genome sequencing and stored in ethanol. Specimens were placed in 70% ethanol for long-term storage. All collections from Lake Masoko were conducted under research permits issued by the Tanzania Commission for Science and Technology.

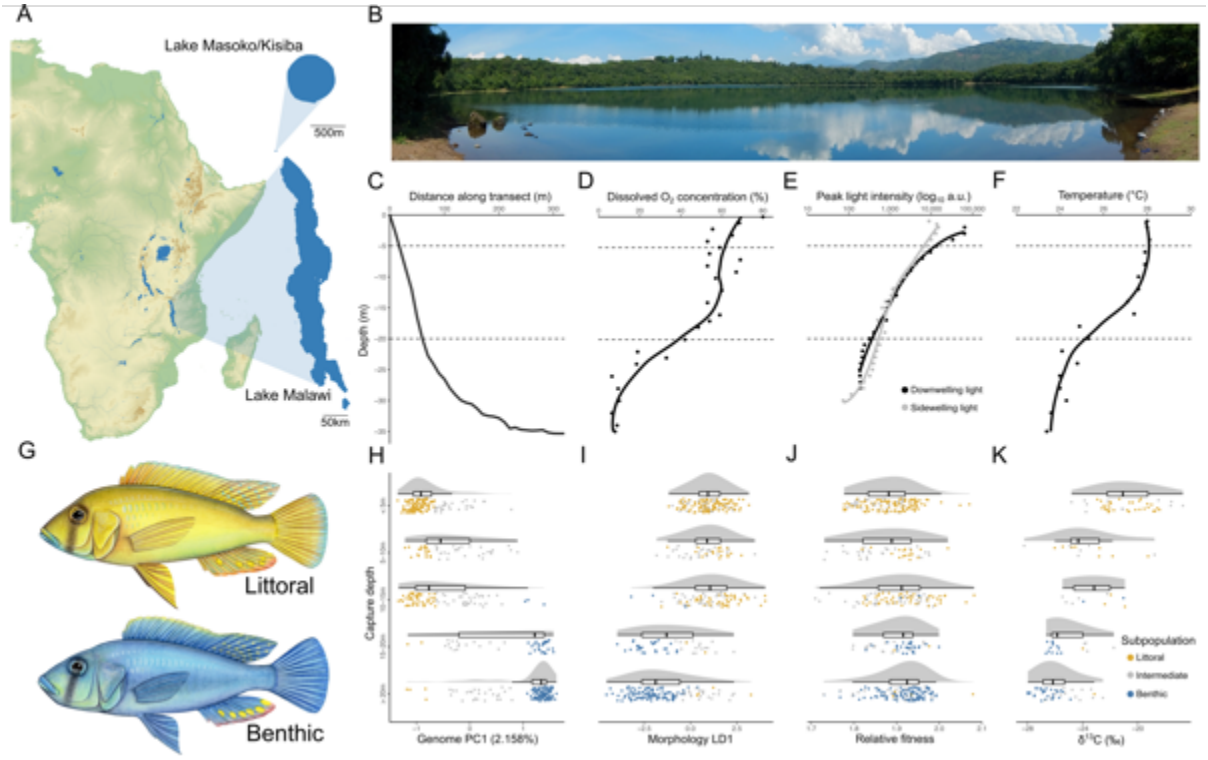


Figure 5.1 Incipient ecological speciation in an East African crater lake. (A) Location of Lake Masoko relative to Lake Malawi and its neighbouring countries, located approximately 30km NW of Lake Malawi's northmost point. (B) Panorama of Lake Masoko from its southern shore. (C) Bathymetric profile of Lake Masoko averaged across four transects (data collected in summer 2018, from Turner et al. 2019). (D) Dissolved O_2 concentration by depth (sampled on 13/03/2005, Delalande 2008). (E) Peak light intensity by depth for downwelling (black) and sidewelling (grey) light. Both measurements taken on 15/08/2018 (F) Water temperature by depth (sampled on 13/03/2005 and 14/03/2005, Delalande 2008). (G) The littoral (yellow, top) and benthic (blue, bottom) ecomorphs of *Astatotilapia calliptera* (paintings by Julie Johnson, lifesciencestudios.com). (H) PC1 of Principal Component Analysis on genetic variation in *A. calliptera* strongly aligns along a depth gradient. (I) LD1 of linear discriminant analysis on 18 morphometric measurements in *A. calliptera*. The benthic ecomorph tends to have a deeper and shorter head. (J) Relative fitness of sampled individuals, estimated using standard length as a proxy. (K) $\delta^{13}C$ ‰ by capture depth, determined by stable isotope analysis of muscle tissue. The benthic fishes have an isotopic signature enriched with ^{13}C, indicating a more offshore zooplanktivorous diet. The littoral fishes have an isotopic signature depleted in ^{13}C, indicating a more inshore diet comprised of littoral invertebrates (Carruthers et al. 2022).

5.4.2 Morphological data

We measured 19 linear morphometric measurements from all 491 individuals following Snoeks (2004) (Figure S5.2), specifically: standard length (SL), body depth (BD), head length (HL), head width (HW), inter-orbital width (IOW), snout length (SNL), lower jaw length (LJL), cheek depth (CHD), premaxillary pedicel length (PPL), eye diameter (ED), lachrymal depth (LAD), dorsal fin base length (DFB), anal fin base length (AFB), predorsal distance (PRD), preanal distance (PRA), prepectoral distance (PRP), prepelvic distance (PRV), caudal peduncle length (CPL) and depth (CPD). The lower pharyngeal jaw (LPJ) was extracted, photographed and weighed for a subset of

306 individuals. For 184 individuals, we measured the width of the LPJ (LPW) and the width of four of the posterior-most teeth. The mean width of these four teeth was calculated (TWA). We counted the number of eggspots on the anal fin of a subset of 346 males.

For the quantification of lateral line system phenotypes, we microCT-scanned a subset of 199 individuals across a range of depth bands. We used a Nikon XTH225ST micro-computed tomography (micro-CT) system (in the University of Bristol), with 3,141 projections and a voxel size of 20-30μm per scan. Following reconstruction of scans we imported stacks of images into VG Studio MAX 3.3.6 (Volume Graphics GmbH, 2016). We reconstructed a 3D model of each scanned fish and captured 2D images of these scans from the ventral head and lateral head perspectives. We used a landmark-based geometric morphometric approach to estimate cephalic canal pore area by drawing curves around each canal pore and resampling to nine sliding semi-landmarks, using the tps suite of software (Rohlf 2015). We used the packages geomorph 4.0.4 (Adams et al. 2022) and geometry 0.4.5 (Habel et al. 2019) in R 3.1.3 (R Core Team 2020) to estimate the area of each pore, by considering each coordinate as a vertex of a polygon, and estimating area using the triangle method. Before further analyses, we estimated the mean pore area for four cranial lateral line canals (mandibular, preopercular, infraorbital, otic) following Edgley and Genner (2019). For a full breakdown of these data, see chapter 2.

All phenotypic traits were $\log_{10}$ transformed, and data for individual traits were standardised using general linear models (predictor variable was $\log_{10}$ standard length; response variables were $\log_{10}$ measured traits). Where morphological trait data were collected by multiple researchers, we accounted for this using researcher identity as a predictor variable. Standardised residuals from these models were used as phenotype variables in all subsequent analyses. All analyses included only males, enabling us to eliminate variation due to sexual dimorphism (Munby et al. 2021), and enabling evaluation of the magnitude of selection on male-specific traits (e.g. eggspots). To visualise gross morphological variation between groups of fish, we conducted a Linear Discriminant Analysis (LDA) including the 18 $\log_{10}$ transformed linear morphometric traits (excluding standard length) using the R package MASS 7.3.57 (Ripley et al. 2022).

5.4.3 Stable isotope analysis

A subset of 128 individuals from a range of capture depths were processed for $\delta^{13}C$ stable isotopes, providing an indication of position of the individual on a diet axis ranging between inshore littoral arthropods to offshore zooplankton (Malinsky et al. 2015; Carruthers et al. 2022). Dried white

muscle samples from the right side of the body were processed by Iso-Analytical (Crewe UK), following methods described in Carruthers et al. (2022).

5.4.4 Whole genome sequencing and variant calling

Whole genome shotgun sequencing data were collected for all 491 Lake Masoko individuals, from preserved pectoral fin clips, using an Illumina HiSeq2000 or XSeq2000. Median coverage was 14.5x, and reads were aligned to the fAstCal1.2 *Astatotilapia calliptera* reference genome (accession GCA_900246225.3). Reads were aligned in bwa (Li 2013) using the mem algorithm (Li and Durbin 2009), and variants were called using GATK 4.2.3 with HaplotypeCaller (McKenna et al. 2010; Poplin et al. 2018). We filtered to only biallelic sites, and conducted quality filtering for sequencing depth, missing genotypes, mapping quality, excess heterozygosity and minor allele frequency. For full detail of sequencing, mapping, variant calling and quality filtering see Munby et al. (2021).

5.4.5 Genomic characterisation of subpopulations

In total 3,328,052 single nucleotide polymorphism sites (SNPs) were used for genomic Principal Component Analysis (PCA) in plink 2.00 (Purcell et al. 2007). The primary axis of genomic differentiation (PC1; 2.158%) was highly correlated with capture depth (Figure 5.1H, Figure 5.2A), and the morphological traits associated with each ecomorph, whereas PC2 (0.33%) separated genetically intermediate individuals from the littoral and benthic individuals (Figure 2A). We categorised individuals as separate ecomorphs or subpopulations genomically according to their genomic PC1 scores (Figure 5.2A). Maintaining consistency of group membership with Munby et al. (2021), "benthic" ecomorphs are PC1 > -0.06, "littoral" individuals are PC1 < 0.0471, and "intermediate" individuals are $-0.06 \leq PC1 \leq 0.0471$.

5.4.6 Fitness proxy

Here we use body size as a proxy for fitness, integrating growth and survivorship. Larger individuals in haplochromine cichlid fishes are likely to have had the greatest reproductive success prior to the time of sampling. This is for three reasons. First, cichlid fish exhibit Type III survivorship curves, so most individuals in a population will not reach maturity (Trendall 1986). Second, the ability to successfully defend breeding territories is closely linked to body size in adult male haplochromines (Markert & Arnegard, 2007), and without a territory a male is unlikely to secure matings (Jordan et al. 2010). Third, size, age and reproductive status are strongly positively correlated in male *A. calliptera* (Figure S5.3A), and therefore fish that reach greater ages and body sizes will on average have a high reproductive output over their lifetime (Figure S5.4). In addition,

this proxy is likely to be a conservative measure of evolutionary fitness, as smaller fit individuals that will go on to grow larger and have significant reproductive success will be considered as being "unfit" in the analysis.

5.4.7 Visualising adaptive landscapes

We visualised the adaptive landscapes within Lake Masoko using GAM smoothing in the package mgcv 1.8.40 (Wood 2011). We smoothed over both genomic space (PC1 and PC2 of genomic PCA) (Figure 5.2A), and discriminant morphospace (LD1 and LD2 of LDA) (Figure 5.2B), using the $\log_{10}$ standard length fitness proxy. We visualised these fitness landscapes with both 2D and 3D plots.

5.4.8 Trait-based selection

For each morphological or ecological trait, we generated linear and quadratic selection gradients using our standard length fitness proxy, following approaches developed by Lande and Arnold (1983). Quadratic selection coefficients were doubled (Stinchcombe et al. 2008). We conducted these analyses for the whole lake, and for each of the three genomic subpopulations in isolation (benthic, littoral, and intermediate). We tested for differences in linear selection between the benthic and littoral subpopulations using a general linear model (GLM) with a gaussian fit, including relative fitness as the response variable, and the interaction between trait values and subpopulation as our predictor. We considered a significant interaction term between subpopulations (benthic and littoral) as indicative that directional selection within subpopulations is leading to disruptive selection between subpopulations. It is worth noting here that the presence of a significant interaction term does not inform about whether selection is stabilising or disruptive, just that it is operating differently between subpopulations. We also estimated selection differentials for each trait (Lande and Arnold, 1983). These were defined as the difference in trait value between the fittest 20% of individuals (according to our fitness proxy), and the mean trait value for all individuals within the subpopulation. We tested for the statistical significance of these differentials using a Monte Carlo randomisation test with 100,000 iterations.

5.4.9 Genome-wide association studies on relative fitness

To test for potential regions of the genome under selection in each ecomorph, we conducted genome wide association studies (GWASs) using $\log_{10}$ standard length as a response variable. After mapping, variant calling and quality filtering, associations between SNPs and $\log_{10}$ standard length were quantified through linear mixed models (LMMs) in GEMMA 0.98.1 (Zhou and Stephens,

2012). We included a pairwise kinship matrix, calculated in GEMMA, in our analysis to account for the effect of shared ancestry. To account for any confounding effects of population stratification (Figs 1H and 2A), PC1 and PC2 of genomic PCA were included as covariates in all linear mixed models (LMMs) given the population structure along these axes (Figure 5.2A) (Zhao et al. 2018).

We conducted GWASs on all male fish in the lake, and also separately for the littoral, intermediate and benthic subpopulations. SNPs which were significantly associated with fitness were identified using a likelihood ratio test ("p_lrt") alongside a Bonferroni-corrected 5% significance threshold. To prevent overcorrection for multiple testing due to the conservative nature of Bonferroni-correction (Johnson et al. 2010), we also considered loci exceeding a 0.5% FDR correction threshold (Kaler and Purcell, 2019) and an indicative 0.00001 threshold.

5.5 Results

5.5.1 Genomic and morphological divergence in Lake Masoko

The genomic fitness landscape within Lake Masoko revealed two distinct peaks at the extreme values of PC1 (the primary axis of genomic differentiation) coinciding with distinct clusters of deep-caught benthic fish and shallow-caught littoral fish (Figure 5.2A & C). Genomically-intermediate individuals occupied a fitness valley, generating a pattern consistent with disruptive selection along the benthic-littoral axis (quadratic selection gradient γ=0.1598; Table 1). A similar bimodal pattern was observed in discriminant morphospace, which captured variation in a suite of 19 morphological traits (Figure 5.2B & D). Extreme values of LD1 (the primary axis of morphological differentiation) showed relatively high level of fitness, and were characterised by distinct clusters of the deep-caught benthic fish and shallow-caught littoral fish. Between these extremes was a fitness valley, occupied by genomic intermediates and some shallow-caught littoral fish, consistent with disruptive selection (quadratic selection gradient γ=0.0181; Table 1). Notably, the fish in the genomically intermediate category tended to have high genomic and phenotypic similarity to littoral fish (Figure 5.1G-H).

5.5.2 Lake wide patterns of selection on individual traits

Across all individuals from the lake, we found no significant directional selection gradient (β) (Figure S5.1A) for any morphological trait (Table 5.1). There was, however, evidence for lake-wide selection against a ^{13}C enriched diet (β = -0.038) (Figure S5.37), characteristic of the fish feeding primarily on littoral arthropods. When calculating quadratic selection gradients (γ) (Figure S2B &

S2C), we found several instances of disruptive and stabilising selection (Table 5.1). In addition to the primary axes of genomic and morphological diversification (Genomic PC1 γ = 0.159816; Morphological LD1 γ = 0.018092) (Figure S5.5), we found incidences of disruptive selection operating on lower jaw length (LJL) (γ = 0.031306) (Figure S5.16) and eggspot number (γ = 0.021168) (Figure S5.35), while stabilising selection was evident for lachrymal depth (LAD) (γ = -0.03506) (Figure S5.20) and eye diameter (ED) (γ = -0.02781) (Figure S5.19) (Table 5.1).

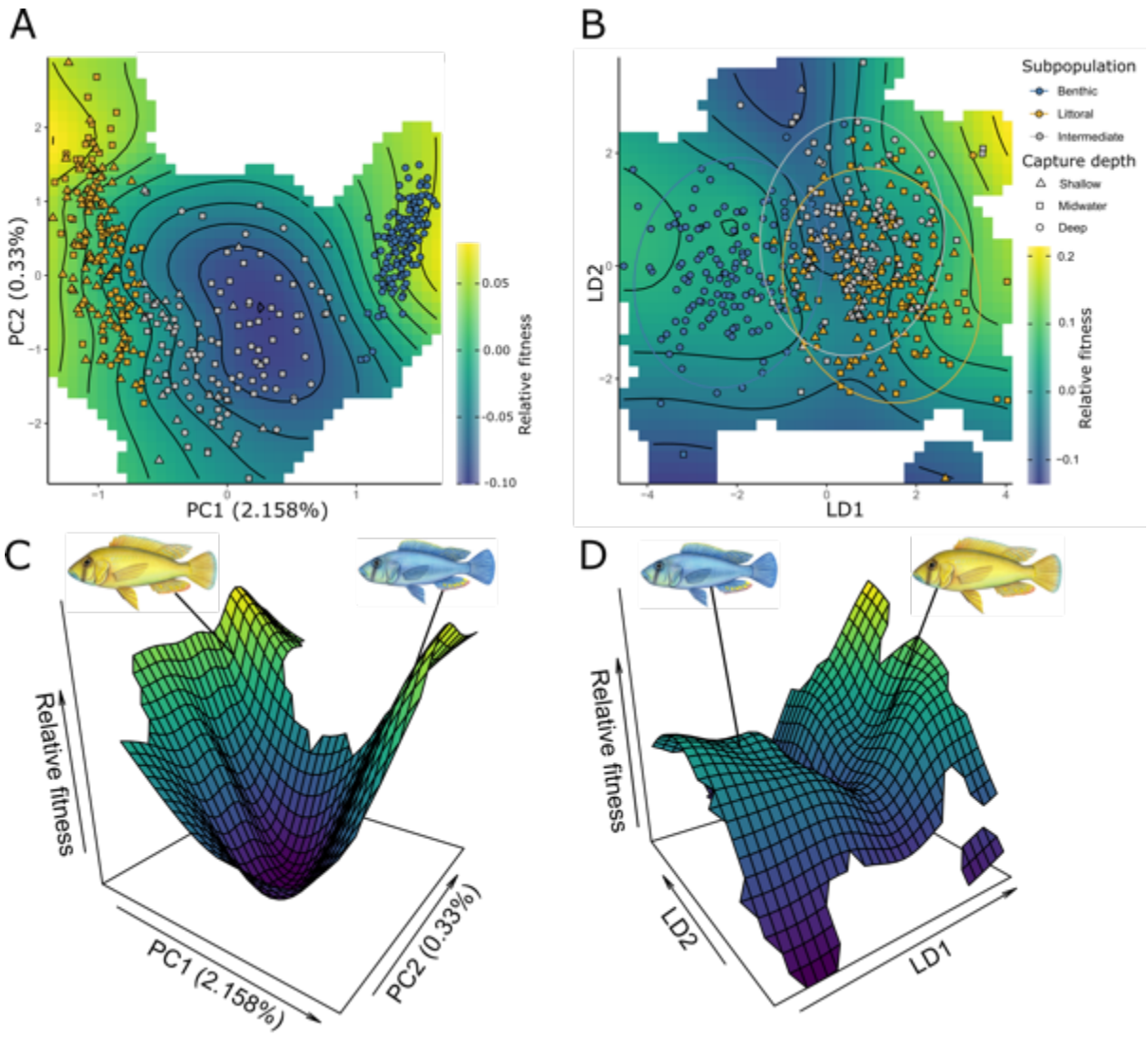


Figure 5.2 Disruptive selection on the genome and body morphology in Lake Masoko. (A) Genomic fitness landscape, using the first two axes of variation across the whole genome captured using a Principal Component Analysis (PCA). (B) Morphological fitness landscape, using the first two axes of variation across 18 linear measurements captured using a Linear Discriminant Analysis (LDA). (C) 3D rendering of the fitness landscape shown in A, (D) 3D rendering of the fitness landscape in B. Relative fitness is estimated using relative standard length as a proxy.

5.5.3 Subpopulation-specific directional selection leads to divergent phenotypes

In contrast to the lake-wide analyses, we found evidence of several traits being under significant directional selection within subpopulations, and the direction of these within-population selection gradients was not always consistent among subpopulations (Table 5.1). Divergent patterns of selection between genetic subpopulations were observed in multiple traits, as indicated by significant GLM interaction terms and viewing contrasting trajectories on fitness surfaces (Figs S5-S37), including lower pharyngeal jaw weight (LPG), average lower pharyngeal jaw tooth width (TWA), premaxillary pedicel length (PPL), dorsal fin base length (DFB), caudal peduncle length (CPL) and diet as measured using $\delta^{13}C$ (Table 5.1). There was no apparent overlap of individual traits in which significant disruptive selection was found using quadratic selection gradient approach, and the approach that explicitly considers subpopulation structure (Table 5.1).

Table 5.1 Quantifying selection in Lake Masoko. Significance of linear directional (β), stabilising (γ<0) and disruptive (γ>0) selection on genomic, morphological, and ecological variables in Lake Masoko. Linear selection gradients are shown for the whole lake, as well as individually for benthic, littoral, and intermediate subpopulations. Also shown is the p value for the GLM interaction term between the benthic and littoral subpopulations. It is worth noting that a significant interaction term does not inform about whether selection is stabilising or disruptive, just that it is operating differently between subpopulations. Significant selection gradients or p-values are highlighted in bold. *p<0.05; **p<0.01; ***p<0.001.

Standardised Trait	Linear selection gradient β (lakewide)	Quadratic selection gradient γ (lakewide)	Linear selection gradient β (littoral)	Linear selection gradient β (intermediate)	Linear selection gradient β (benthic)	GLM benthic-littoral interaction term p-value
Genome PC1	0.009188	**0.159816*** **	**-0.19603** **	**-0.11165*** **	0.07646	**0.0003503*** **
Genome PC2	**0.032080*** **	0.001322	0.01673	-0.001376	0.004859	0.1467
Morphology LD1	-0.0001352	**0.018092*** **	**0.03079** **	**0.06845*** **	0.01143	**5.041e-11*** **
Morphology LD2	-9.445e-05	-0.01214	0.01750	0.009745	0.012943	0.1322
Body shape PC1	-0.011709	0.014736	-0.04126	**-0.04306***	**-0.04435** **	0.5156
Body shape PC2	0.008249	0.019752	-0.0008017	0.01080	0.01317	0.5382
Body depth	9.646e-06	-0.00034	-0.0009909	**0.02863***	**-0.024250***	0.1682
Head length	-6.837e-05	0.008651	**-0.02746***	-0.03305	**-0.03362***	0.4502
Head width	-4.525e-05	-0.01965	-0.006564	0.006954	-0.007510	0.7294
Interorbital width	-4.255e-05	-0.01291	0.004504	0.01099	-0.001758	0.6798
Snout length	-6.898e-05	0.000194	-0.001044	-0.005997	-0.017884	0.6993
Lower jaw length	-3.947e-05	**0.031306** **	-0.02193	**-0.05135** **	-0.02802	0.6323
Premaxillary pedicel length	-5.363e-05	0.016732	**-0.03825** **	-0.03408	-0.01039	**0.03676***
Cheek depth	-4.899e-06	-0.00551	-0.01239	-0.008523	-0.01900	0.6871
Eye diameter	1.085e-05	**-0.02781** **	-0.001883	-0.01763	0.019293	0.1664
Lachrymal depth	-6.181e-05	**-0.03506*** **	0.006932	-0.01196	-0.01491	0.4857
Dorsal fin base length	0.000078	0.012732	**0.03379***	**0.06000*** **	**-0.02748***	**0.0006127*** **
Anal fin base length	-9.971e-05	0.000419	-0.001524	0.0006164	0.001576	0.859
Predorsal distance	-1.168e-05	-0.00085	-0.01080	-0.01775	-0.009652	0.7313
Preanal distance	2.557e-05	0.000451	-0.008165	-0.0007015	-0.002291	0.569
Prepectoral distance	-0.0000836	-0.00383	-0.002642	**-0.05148** **	**-0.04177** **	0.7365
Prepelvic distance	-6.822e-05	-0.00634	0.00461	-0.009646	**-0.04187*** **	0.105
Caudal peduncle length	2.093e-05	0.01175	-0.01249	0.007365	**0.036325*** **	**0.01063***
Caudal peduncle depth	2.113e-05	-0.00069	-0.004919	**0.03808** **	-0.007912	0.4292
LPJ mass	3.169e-05	0.018808	**0.04213***	0.0008091	0.06733	**0.03845***
LPJ width	-0.0001704	0.0147	0.008892	0.01155	-0.02366	0.4183

Average LPJ tooth width	0.0000986	0.010598	0.03900	0.01661	-0.03840	**0.02752***
Mandibular canal	--0.0003472	-0.00198	0.01357	0.001714	**-0.03190***	0.1277
Preopercular canal	-0.000279	0.015178	-0.01082	-0.03486	-0.006321	0.7934
Infraorbital canal	-0.0005458	0.003363	-0.02453	-0.04056	-0.02567	0.808
Otic canal	-0.000329	-0.0061	-0.005841	-0.02149	-0.01577	0.9956
Eggspot Number	2.684e-05	**0.021168***	0.01604	0.01458	0.005208	0.7822
Diet ($\delta^{13}C‰$)	**-0.03828*****	-0.03272	-0.04914	**-0.08770*****	**0.10081****	**0.0196***

5.5.4 Comparing signatures of present-day selection with average trait values in subpopulations

To determine whether current patterns of selection consistently reflect the pattern of past evolutionary change, we compared the magnitudes and directions of our measured selection differentials with the differences in trait magnitudes between the two ecomorphs. In many cases, the direction of selection differentials (Figure 5.3A) was the same as the mean trait value differences between ecomorphs (Figure 5.3B), including for genomic PC1, body depth, lower jaw length and dorsal fin base length. However, this pattern did not hold in all cases. For example, shorter premaxillary pedicels appear to be favoured by selection in the benthic population, yet on average benthic fish have longer premaxillary pedicels than littoral fish (Figure 5.3).

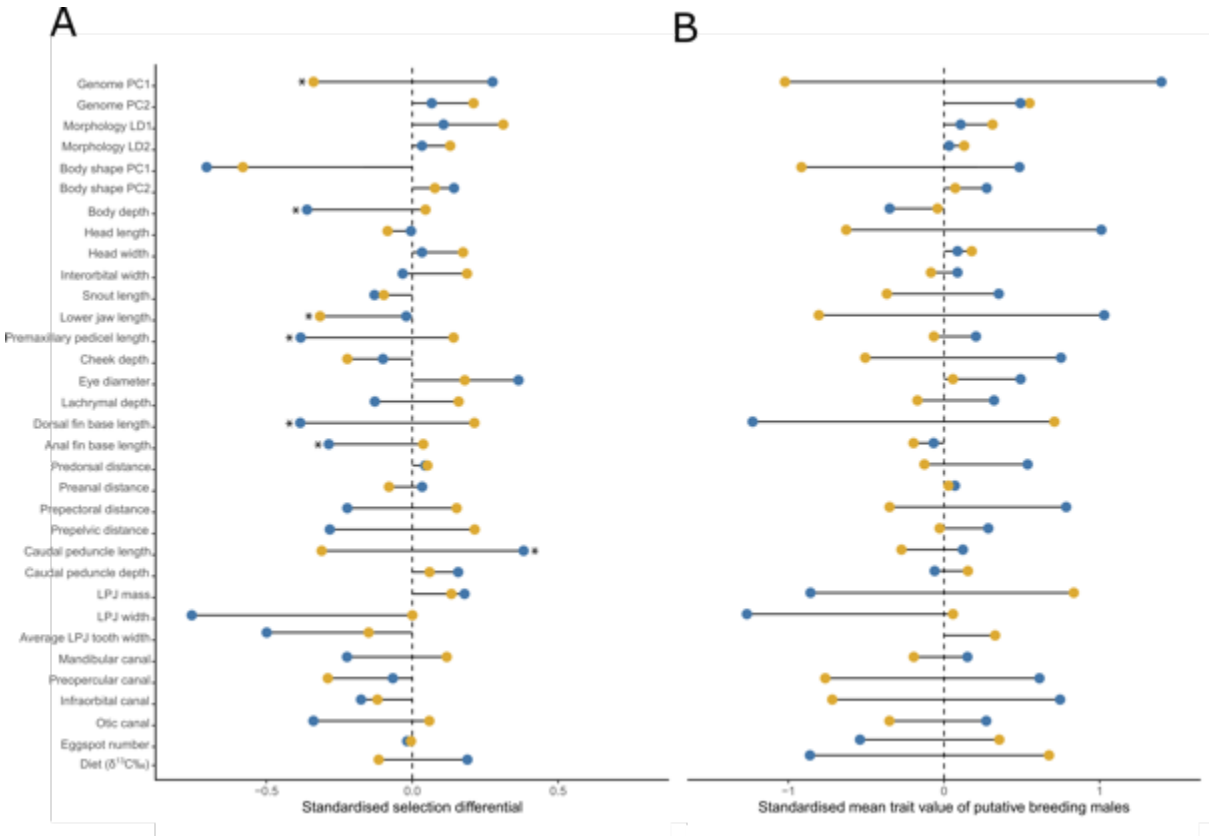


Figure 5.3 Current patterns of selection in relation to past evolutionary change in Lake Masoko. (A) The standardised selection differential for each trait within subpopulations. Selection differentials are defined as the difference in standardised mean trait value for the fittest 20% of individuals and the mean trait value for the subpopulation. *Significant selection differentials (only shown for traits in which sample size was sufficient. This excludes gross body shape and LDA measures, LPJ

measurements, lateral line system measurements and diet, full details in Table S5.2). (B) The standardised trait value for each trait in the fittest 20% of males, within each subpopulation.

5.5.5 Subpopulation-specific disruptive selection on the genome

We used linear mixed models in GEMMA to test for associations between genomic variants and our proxy fitness measure. Across the 3,328,052 SNPs in our callset, we identified a set of 54 SNPs that crossed the P<0.00001 threshold in the littoral population, and 15 that crossed the P<0.00001 threshold in the benthic population (Figure 5.4B-C). Using the more conservative 0.1 FDR threshold, we only identified 11 SNPs in the benthic population, and two that crossed this threshold in the benthic population. These contained SNPs proximate to genes associated with a range of functions from immune system function to growth and metabolism, with limited overlap with existing evidence of genes under selection (Malinsky et al., 2015; Carruthers et al., 2022; Vernaz et al., 2022). The regression beta coefficients of individuals SNPs can be interpreted as linear selection gradients. These show an overall pattern of consistent directions of associations with fitness within each of the two populations (Figure 5.4D). Notably, we found that loci with the strongest statistical signals of association with fitness (which cross the P<0.00001 threshold,) were unique to only one of the two populations. Thus, similarly to individual morphological traits, individual SNPs that most strongly under directional selection in benthic subpopulation were not necessarily under significant directional selection in the littoral subpopulation and vice versa.

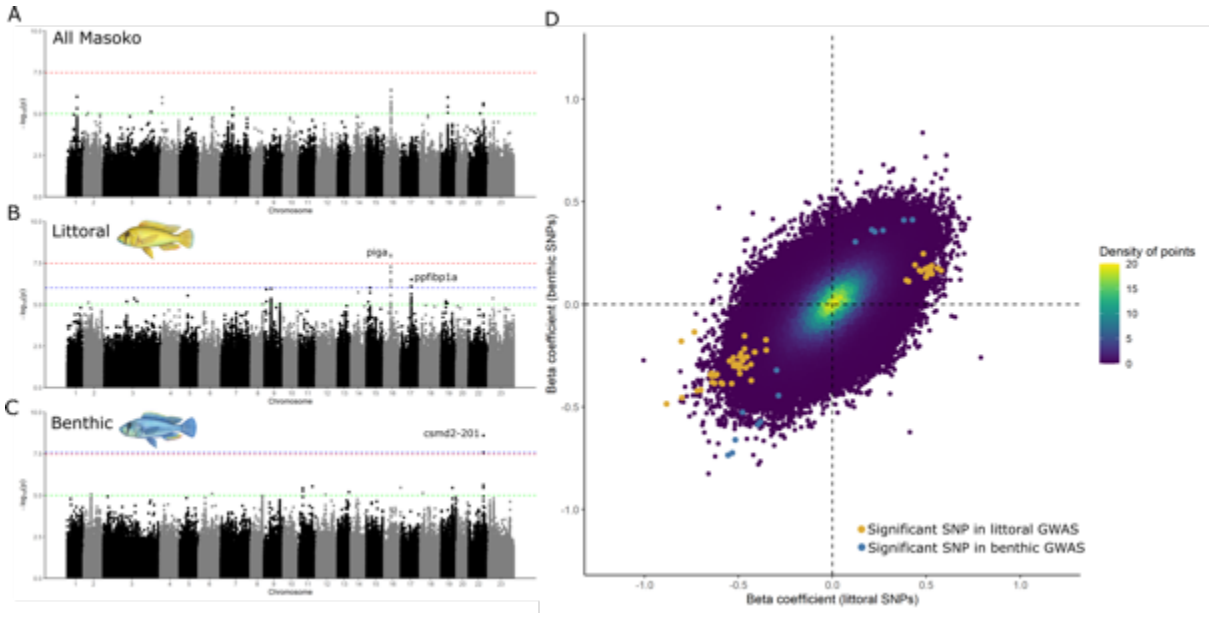


Figure 5.4 Genome-wide association study of relative fitness in Lake Masoko. (A-C) Manhattan plot SNPs associations with fitness lakewide, in littoral fish and benthic fish, respectively. Red line is Bonferroni threshold; Blue line is 0.1 FDR correction threshold; Green line is P<0.00001 threshold. (D) Beta regression coefficients of SNPs in the littoral GWAS (x-axis)

and benthic GWAS (y-axis). Yellow and blue points are SNPS below a P<0.00001 threshold in the littoral and benthic fish respectively.

5.6 Discussion

Our findings here reinforce evidence for the existence of multiple ecologically, phenotypically and genomically distinct subpopulations of *A. calliptera* in Lake Masoko (Malinsky et al. 2015; Carruthers et al. 2022; Vernaz et al. 2022). In addition, we find evidence for disruptive selection on both the genome and gross morphology between these subpopulations, with the benthic and littoral subpopulations occupying independent fitness peaks on genomic (Figure 5.2C) and morphological (Figure 5.2D) adaptive landscapes. We also investigated the benefit of studying disruptive selection on individual traits at the outset of population divergence, by considering subpopulations in isolation, rather than the classic approach of considering species or populations in their totality (Lande & Arnold, 1983).

5.6.1 Evidence for selection against intermediate individuals

When characterising our genomic subpopulations, we find evidence of a genomically and phenotypically transitional group of cichlids, with a broad distribution across depth zones (Figure 5.2). When quantifying the relationships between fitness and their overall genomic composition, and between fitness overall body shape, we find strong evidence that they have comparatively low fitness (Table 5.1). This was evident from fitness landscapes, where genomically intermediate individuals occupied clear fitness valleys between littoral and benthic fitness optima (Figure 5.2C-D) (Coyne and Orr, 2004). In addition, we found evidence for strong positive directional selection (linear selection gradient $\beta = 0.2361$) along the second genomic axis (PC2; 0.33%) (Table 5.1), indicating selection against low values of Genomic PC2, where phenotypically and genetically intermediate individuals are clustered.

It may be the case that these "intermediates" are merely benthic-littoral hybrids, and are hence poorly adapted to the requirements of both ecomorphs' respective habitats, resulting in low relative fitness. However, the existence of apparent population structure along the second axis of genomic differentiation (Genome PC2) (Figure 5.2A) may suggest that these intermediate individuals are not transient hybrids, but a somewhat persistent (albeit low fitness) population, with a degree of reproductive isolation from both benthic and littoral ecomorphs. Intermediate fish are phenotypically extremely similar to the littoral subpopulation in most traits (Supplementary Figure S5.5 – S5.37), and they overlap substantially in discriminant morphospace with littoral fish (Figure 5.2B). As such, rather than hybrids, these fish may simply be a subset of the least fit littoral individuals. If this is the case, their apparent population structure may reflect geographic isolation of less fit individuals by competitive exclusion (Winkelmann et al., 2014), leaving these least fit

smaller individuals to breed amongst themselves. Further information on the ecology, life history, and degree of gene flow between these subpopulations is needed to determine the role of this intermediate subpopulation in the dynamics of ecomorph divergence in Lake Masoko.

5.6.2 Measuring disruptive selection

Conventionally, we would expect the signal of disruptive selection on a trait to manifest as a significant quadratic selection gradient (Lande and Arnold 1983) and multiple peaks on a fitness surface (Kingsolver et al. 2001) or adaptive landscape (Martin and Wainwright 2013) (Figure S5.1). These approaches have become the conventional methods for inferring disruptive selection in wild populations (Kingsolver et al. 2001; Martin and Wainwright 2013; van Rijssel et al. 2018; Moser et al., 2018). However, in parapatric systems where there is environmental heterogeneity within the population, such as along a cline, we may expect to observe distinct patterns of linear directional selection within microhabitats. The Lake Masoko system is one such example of this (Malinsky et al. 2015), and evidence is mounting that ecological speciation can occur, and may even be common, in sympatry along these environmental gradients (Lande 1982; Doebeli and Dieckmann 2003; Rundle and Nosil 2005; Sulloway and Kleindorfer 2013; Malinsky et al., 2015). In these systems, even subtle differences in magnitudes of directional selection will lead to contrasting phenotypes over time, even if the direction of selection is the same between the two. For example, in a system where strong positive linear selection operating on one microhabitat, but weak positive linear selection is operating in the other, the differences in the rate of change will over time lead to phenotypic divergence (Figure 5.3B). This kind of disruptive selection may not be captured by the conventional method of calculating quadratic or higher order selection gradients (Figure S5.1), which accounts for neither population genetic structure nor differential microhabitat use, both of which may be affecting phenotype's relationship with fitness.

In addition, theoretical models of sympatric speciation often assume that selection will promote a bi-directional pattern of evolution, such as would cause a significant positive quadratic relationship on a fitness surface (Lande and Arnold 1983) (Figure S1C). This may be the case in specific circumstances, and disruptive selection gradients may be powerful at detecting incidences of rapid bidirectional divergence in island systems such as Darwin's finches (Hendry et al. 2009). However, given that it has become increasingly clear that other models of speciation with gene flow may be common (Coyne and Orr, 2004; Rundle and Nosil, 2005; Barluenga et al. 2005; Nosil 2005; Malinsky et al. 2015), selection pressures in these circumstances may not always be driving truly bidirectional phenotypic change. In peripatric systems, for example, it is likely that the

subpopulation moving into a new niche will experience directional selection on traits, in contrast with the parent population, which does not. It may experience selective sweeps, altering the range of viable phenotypes within this semi-isolated population. In parapatry, different ecological circumstances along the environmental cline will necessitate contrasting fitness patterns, with traits being more tightly or loosely tied to fitness in different habitats. In all these circumstances, disruptive selection may not be manifesting as multiple fitness peaks, but subtler patterns in the fitness surface.

Moreover, when significant quadratic selection gradients are found in wild populations, they are generally weak, and despite the fact that very strong directional selection can occasionally occur (Hoekstra et al. 2001), strong disruptive selection has never been observed (Kingsolver et al. 2001). This absence of incidences of strong disruptive selection may indicate that divergence often occurs through these subtle differences in fitness relationships between subpopulations or microhabitats, much of which is undetectable using quadratic selection gradients. When exploring patterns of disruptive selection within two diversifying subpopulations, we identified three forms presently operating on traits within Lake Masoko (Figure 5.5).

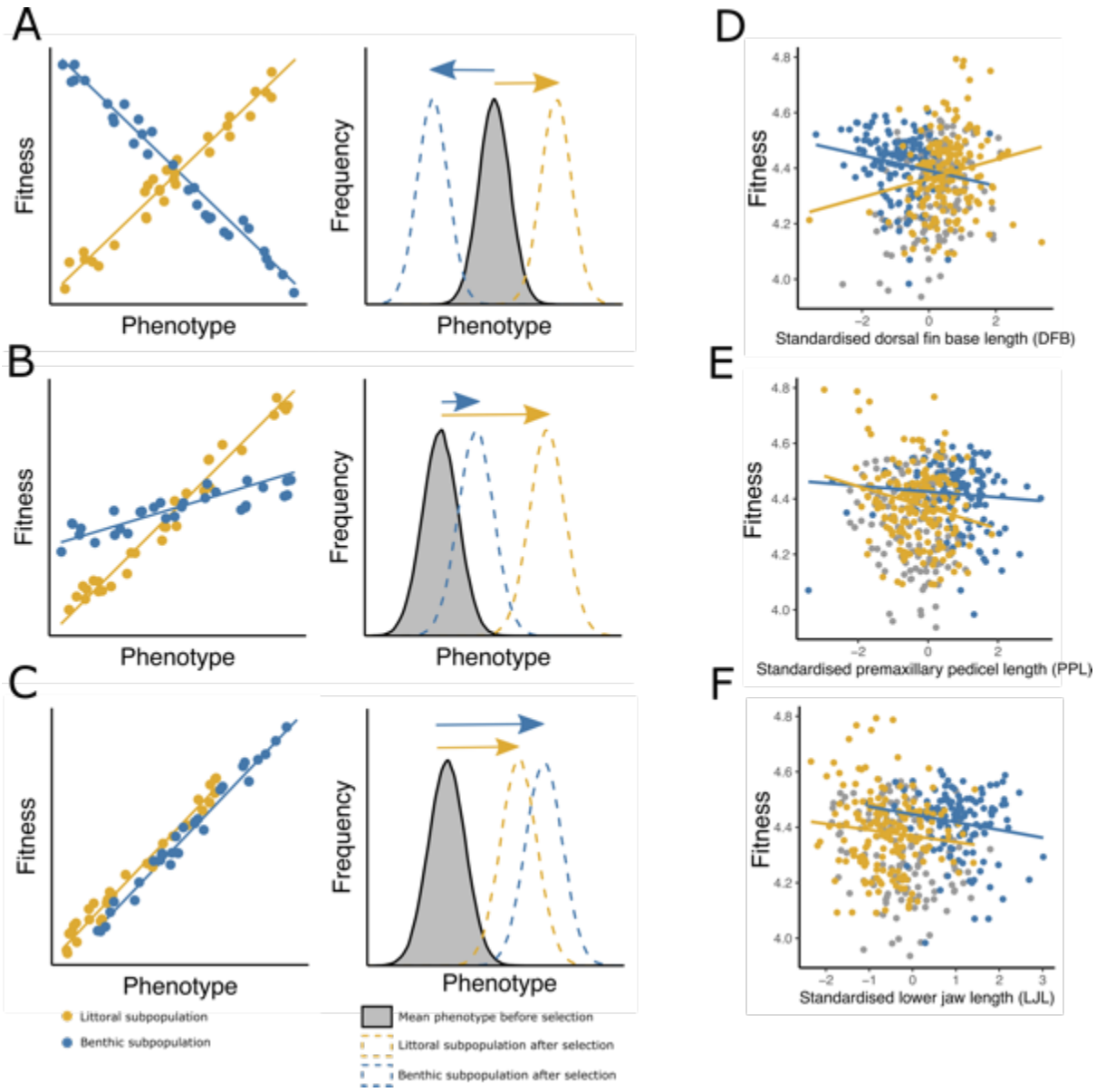


Figure 5.5 Models of divergence. (A-C) Models for disruptive selection within a system with two subpopulations. Left is a hypothetical relationship between phenotype and fitness within each subpopulation, and right is an induced shift in the phenotype. (A) Type I disruptive selection. Two subpopulations are undergoing significant linear selection for the trait, but one is positive and the other negative, leading to trait divergence. (B) Type II disruptive selection. Two subpopulations are under the same direction of selection, but there is stronger selection in one (yellow) than the other (blue). Alternatively, one subpopulation is under no selection for the trait, but the other is under linear directional selection. These patterns of selection lead to trait divergence. (C) Type III disruptive selection. Two subpopulations are under the same strength and direction of selection, but mean phenotypes differ in their range, leading to consistent differences in the trait between subpopulations. (D-F) Example fitness surfaces for three morphological traits in Lake Masoko *A. calliptera* under the three examples of disruptive selection: (D) type I, standardised dorsal fin base length, (E) type II, standardised premaxillary pedicel length, (F) type III, standardised lower jaw length.

Type I disruptive selection is the scenario where trajectories of directional selection are operating in opposing directions in subpopulations. This this will promote a bidirectional evolutionary

change, with the magnitude of the trait moving towards two extremes in both subpopulations simultaneously (Figure 5.5A-B). This is analogous to the quadratic selection gradient, as defined by Lande and Arnold (1983). In our dataset, we observed Type I selection operating on dorsal fin base length (Figure S21C) and caudal peduncle length (Figure S26C), amongst other phenotypes.

Type II disruptive selection involves directional selection operating simultaneously and in the same direction in subpopulations, but with different magnitudes in each, as evidenced by strongly contrasting selection gradient slopes (Figure 5.5C-D). This will lead to directional phenotypic change in each subpopulation, but ultimately will generate contrasting phenotypes between the two, despite trajectories of change being similar (Figure 5.5D). We observed Type II selection operating on prepectoral distance (Figure S24C) and premaxillary pedicel length (Figure S17C).

Type III disruptive selection is the scenario where directional selection is operating simultaneously and in the same direction in subpopulations, with similar magnitudes as evidenced by similar selection gradient slopes (Figure 5.5E-F). However, despite a similar selection gradient, the phenotypic range of the populations differ, perhaps due to developmental constraints or an absence of appropriate standing genetic variation. We observed Type III selection operating on head length (Figure S12C), head width (Figure S13C), and cheek depth (Figure S18C), among others.

5.6.3 Comparing ongoing disruptive selection to average population phenotypes

We observe scenarios in which the direction and trajectory of linear directional selection is the same in both subpopulations, but there are significant differences between the mean trait values between groups (Figure 5.5E-F), here named type III disruptive selection. It may seem evident that differences in 'adaptive' traits between populations will always be the result of past or ongoing disruptive selection. However, in these cases, the difference in phenotype magnitude between the two may be the result of historic disruptive selection, and the current patterns of selection may not reflect past evolutionary change. In addition, if past disruptive selection is no longer operating, why do the two populations still experience significant directional selection? It may be the case that one subpopulation is developmentally constrained to a different range of phenotypes to another subpopulation, limiting its magnitude, despite the phenotypic extreme being favoured for survival by the same degree in both. This constraint may be related to life-history or ecology, conferred by the subpopulation's respective microhabitat or niche. Alternatively, it may be simply a lack of sufficient genetic variation in one subpopulation constraining the range of possible trait

values. The benthic subpopulation within Lake Masoko has undergone selective sweeps reducing heterozygosity across regions of the genome due to genetic/divergence hitchhiking (Malinsky et al. 2015), which may be constraining the range of possible phenotypes within the deeper water.

5.6.4 Genome wide association-mapping of relative fitness

Our within-subpopulation fitness GWASs revealed clearly contrasting patterns of selection in the two ecomorphs. This shows that the distinct selection pressures operating within the habitats of the two ecomorphs are manifesting on genes differently, which is congruent with previous findings using different approaches to measuring selection on the genome (Malinsky et al. 2015; Carruthers et al. 2022). Using fitness as a trait in methods testing for phenotype-genotype interactions has in the past been limited by constraints on genome sequencing technology and cost. However, in recent years, it has been used successfully to identify candidate genes under selection in QTL mapping in sticklebacks (Schluter et al. 2021), and in GWAS in Cyprinodon pupfishes (Patton et al. 2022). Our findings here provide evidence that different regions of the genome are under selection in the two ecomorphs, implying different genes are important for survival in each. Despite this, these findings are only as accurate as our fitness proxy (Viblanc et al. 2022). Larger sample sizes, and higher sequencing depth is required to obtain the fine-scale resolution needed for more accurate phenotype-genotype mapping on traits as broad as a proxy for relative fitness (Patton et al. 2022).

5.6.5 Validation of fitness proxy

The validation of these results is highly dependent on the accuracy of our choice of fitness proxy (Hoekstra et al. 2001; Franklin and Morrissey 2017; Viblanc et al. 2022). In haplochromine cichlids, growth is indeterminant, and age is tightly linked with size under controlled conditions (Tweddle and Turner 1977). We found that *A. calliptera* individuals increased in size linearly over time under controlled conditions in aquaria (Figure S5.3A). Given enough space and food, they will tend to continue to grow at a relatively constant rate. In addition, body size is tightly linked to reproductive success in haplochromine cichlids, because larger males are better able to hold territories and will be the only individuals to spawn (Markert & Arnegard 2007) and sneaky matings are not known in haplochromine cichlids such as *A. calliptera*. We found that larger *A. calliptera* individuals had larger gonads as a proportion of total body weight, within both males (Figure S5.3A & S5.3C) and females (Figure S5.3B & S5.3D), which is consistent with existing evidence that larger fish have disproportionally increased reproductive output (Barneche et al. 2018). It could be argued that any individual smaller fish may go onto grow into a larger fish and go on to have lots of reproductive

success. However, given the relatively consistent growth rate, on average a younger and smaller fish will, in most cases, have fewer offspring than a larger fish that has already beaten the odds of being predated during the juvenile and sub-adult stages. In this sense, body size is a useful proxy for survivorship in a predator-rich environment such as Lake Masoko, with precedent for use from other studies (Martin and Pfennig 2009; Franklin and Morrissey 2017).

Despite this, confirmation of the reliability of this fitness proxy, or the identification of a better alternative for wild-caught individuals, is required. This may be established by exploring the genomic contributions of individuals to offspring sampled in the wild, or by experimental approaches such as field mesocosms, or even observational work scoring mating success in the natural habitats. For all individuals, we also followed an established approach involving counting growth rings (circuli) on scales, to use as a proxy for age (Doyle et al. 1987; Kingsford and Atkinson 1994) and regressing this against size, to obtain a proxy for growth rate (van Rijssel et al. 2018) (Chapter 5 Supplementary material). Despite finding that true age, number of circuli, and standard length were all correlated (Figure S5.4A-C) (van Rijssel et al. 2018), we did not find that this growth proxy accurately correlated with the true growth rate of individuals (Figure S5.4D), contrary to existing evidence. As such, we conclude that this growth rate proxy was no more accurate than simply using standard length as a measure of individual fitness, and does not encompass survivorship as straightforwardly. The results of our analyses using the 'growth rate' proxy, and the results of using simply standard length were extremely similar, both in terms of adaptive landscapes and selection gradients (data not shown).

5.6.6 Conclusions

Here, we have shown that population stratification within a system, either in terms of microhabitat use or genetic structure, can mask adaptive divergence on traits within the ecological speciation process. We propose models for measuring within-subpopulation fitness surfaces, revealing subtle and complex patterns of evolutionary change operating at the outset of population divergence. We suggest that any indications of stratification within populations, whether it be genetic or ecological, should be considered when measuring selection. The approaches suggested by Lande and Arnold (1983) for measuring quadratic selection gradient may be good approximations for general phenotypic change, but begin to break down in non-panmictic populations across heterogeneous environments. The use of quadratic selection gradients will be particularly limited in situations more complex than Lake Masoko, such as multispecies adaptive radiations, where many

ecologically-divergent populations emerge with contrasting patterns of selection over short periods of evolutionary time.

5.7 Supplementary Information

Supplementary Methods

Ecological data collection

Bathymetric data (Figure 5.1C) were collected through four transects across the lake, which were subsequently averaged. Data on dissolved oxygen concentration and water temperature within depth bands (Figure 5.1D and 5.1F) are from Delalande (2008). Sidewelling and downwelling light intensity data were captured at 2m depth bands on the same day (15th August 2018), using a calibrated Ocean Optics USB200+VIS-NIR-ES spectrometer (Figure 5.1E).

Standard length as fitness proxy

As the accuracy of any measurements of selection gradients is highly dependent on the validity of the proxy fitness measurement (Hoekstra et al. 2001; Viblanc et al. 2022), we conducted preliminary tests to determine which is the most appropriate. Previous studies have used a proxy growth rate as a condition-based proxy for fitness in studies quantifying natural selection, including in related species such as Lake Victoria cichlids (Moser et al. 2018; van Rijssel et al. 2018). These methods involve counting fish scale circuli. Circuli are concentric rings produced on scales at a relatively consistent rate as a fish grows (Viertler et al. 2021; Kingsford and Atkinson 1994). Circulus number is known to be significantly correlated with age, as verified by laboratory growth experiments (Kingsford and Atkinson 1994; Moser et al. 2018; van Rijssel et al. 2018), and is a method commonly used to age fishes (Kingsford and Atkinson 1994; Ilieş et al. 2014; Thomas et al. 2019).

We counted scale circuli on all 491 individuals caught at Lake Masoko. Three non-lateral line scales were collected and mounted from each of the 491 individuals. Due to variation in the morphology of scales based on their position on the fish (Lippitsch 1990; Viertler et al. 2021), we chose only the three anterior-most scales in the first row ventral to the trunk lateral line. Scales were removed with tweezers, cleaned briefly with bleach to remove any residual tissue, and rinsed with water. Scales were mounted on slides, and images captured under a Leica m205-C stereo dissection microscope with a mounted GXCAM HiChrome-Met-M camera and screen. The number of circuli was counted on each, from the outermost ring to the centre of the scale. As scales may be removed and regrown during the lifespan of a fish (Ghods et al. 2020; Yasuaki et al. 1989), we used the highest measurement from the three scale ring counts of each individual for further analyses. To estimate individual fish growth rate, we conducted a linear regression of $\log_{10}$-transformed standard length against circuli number (age proxy), following an established approach

(Kingsford and Atkinson 1994; Doyle et al. 1987; van Rijssel et al. 2018). Residuals from this regression were tested as a potential proxy for growth rate, and hence fitness.

To determine whether this proxy measure for growth rate is an accurate approximation for true growth rate (van Rijssel et al. 2018), we tested if scale circuli are a good approximation for age. We raised five broods of *A. calliptera* and counted their scale circuli according the same method as the Lake Masoko cichlids outlined above. For these 73 individuals of known age, we were able to compare the true growth rate to our proxy for growth rate. For each individual we extracted and mounted 3 non-lateral line scales and counted scale circuli, as outlined above. All five broods were bred from wild caught individuals from five locations in the Lake Malawi catchment (two broods from Chisumulu island; one from Mpatsonjoka dambo; one from the Ruvuma river in Tanzania; and one from Linthipe river in Malawi) (Table S5.1).

Within these individuals of known age, we found that true age, number of circuli, and standard length were all extremely strongly correlated (Figure S5.3A-C). However, unlike as has been previously found (van Rijssel et al. 2018), we did not find that this growth proxy accurately correlated with the true growth rate of individuals, as calculated using age, despite age and circulus number being correlated with each other (Figure S5.3D) (GLM $F_{(55,56)}$=-0.077355, p=0.571). Other research also finds that the number of circuli does not necessarily increase linearly with time, but can increase or slow dependent on temperature or feeding regime (Thomas et al. 2019), similarly to standard length. As such, we conclude that our growth rate proxy is no more accurate than simply using standard length as a measure of individual fitness, but through using standard length, we incorporate aspects of both growth rate and survival into our fitness proxy. The results of our analyses using the scale-determined growth rate proxy, and the results of using simply $\log_{10}$ transformed standard length, were extremely similar, in terms of adaptive landscapes and selection gradients. As such, for all subsequent analyses $\log_{10}$ standard length is used as our proxy for relative fitness.

Gonad weight measurement

In addition, for each individual included in our growth rate validation dataset (Table S5.1), we dissected and weighed gonad tissue. Gonad weight, and gonadosomatic index have previously been used as proxy measures of reproductive condition, or reproductive success, and may be connected to individual fitness (Zeyl et al. 2014). As such, testing for associations between relative gonad weight and our fitness proxies may help to validate them as measures of fitness.

Prior to dissection, each fish's total body mass was weighed on a microbalance. Each individual was dissected and was sex was determined where possible (where specimens were sufficiently mature). All gonad tissue was removed, and its mass was immediately weighed. For each fish we measured gonadosomatic index (GSI), the ratio between weight of gonads and total tissue weight (Flores et al. 2019).

$$GSI = \frac{gonad\ weight}{total\ tissue\ weight} \times 100$$

Table S5.1. Brood information for the growth-rate and gonad weight dataset. All individuals were bred in aquaria In Hull, and source population indicates where the wild-caught parents of the brood were sourced from. All sources are within Lake Malawi or rivers within its catchment. *Individuals unsexable due to being insufficiently mature.

Brood	Source population	Total number of individuals	Number of males	Number of females	Release date of brood	Age (days)	Mean number of circuli
ACZ-24	Chisumulu island	15	NA*	NA*	04/03/2021	142	39.20
ACZ-22	Chisumulu island	15	10	5	11/11/2020	255	65.00
ACD-26	Mpatsonjoka dambo	18	10	8	08/04/2020	472	73.89
ACR-28	Ruvuma river	17	11	6	25/11/2019	607	64.89
ACL-15	Linthipe river	8	4	4	02/06/2018	1147	92.25

Table S5.2. Standardised selection differentials for each trait, for the benthic and littoral subpopulations. Differentials are calculated as the difference in standardised mean trait value between the fittest 20% of individuals, and the mean trait value for the whole subpopulation. P-values are from Monte Carlo randomisation tests run for 100,000 iterations. Shown are differentials for traits for which the sample size was comparable.

Standardised trait	Selection differentials (littoral)	p-value	Selection differentials (benthic)	p-value
Genome PC1	-0.338455477	**0.011517***	0.275324976	0.053441
Genome PC2	0.210151653	0.079264	0.066912152	0.354143
Morphology LD1	-	-	-	-
Morphology LD2	-	-	-	-
Body shape PC1	-	-	-	-
Body shape PC2	-	-	-	-
Body depth	0.045488986	0.620619	-0.360635533	**0.017022***
Head length	-0.084456161	0.287639	-0.004416363	0.486928
Head width	0.174114587	0.122015	0.033612431	0.422241
Interorbital width	0.187732627	0.104725	-0.032709426	0.571738
Snout length	-0.097504089	0.25827	-0.129480824	0.226824
Lower jaw length	-0.316099894	**0.016244***	-0.020675635	0.451773
Premxillary pedicel length	0.141649764	0.172564	-0.382479377	**0.011814***
Cheek depth	-0.222424297	0.069439	-0.100248449	0.72593
Eye diameter	0.180047695	0.113984	0.364228591	0.985116
Lachrymal depth	0.158998274	0.144598	-0.128262844	0.772046
Dorsal fin base length	0.213043517	0.075556	-0.383807751	**0.012382***
Anal fin base length	0.037761285	0.396463	-0.286150414	**0.046935***
Predorsal distance	0.052125279	0.364342	0.044319224	0.600277
Preanal distance	-0.079698635	0.299607	0.033950272	0.423888
Prepectoral distance	0.152151036	0.155167	-0.222540748	0.097378
Prepelvic distance	0.213943707	0.077099	-0.282520346	0.051155
Caudal peduncle length	-0.310462429	0.017805	0.381675207	**0.011726***
Caudal peduncle depth	0.059399055	0.652306	0.157384729	0.179882
LPJ mass	-	-	-	-
LPJ width	-	-	-	-
Average LPJ tooth width	-	-	-	-
Mandibular canal	-	-	-	-
Preopercular canal	-	-	-	-
Infraorbital canal	-	-	-	-
Otic canal	-	-	-	-
Eggspot number	-0.00448604	0.508588	-0.017391266	0.533671
Eggspot contrast	-	-	-	-
Diet ($\delta^{13}C‰$)	-	-	-	-

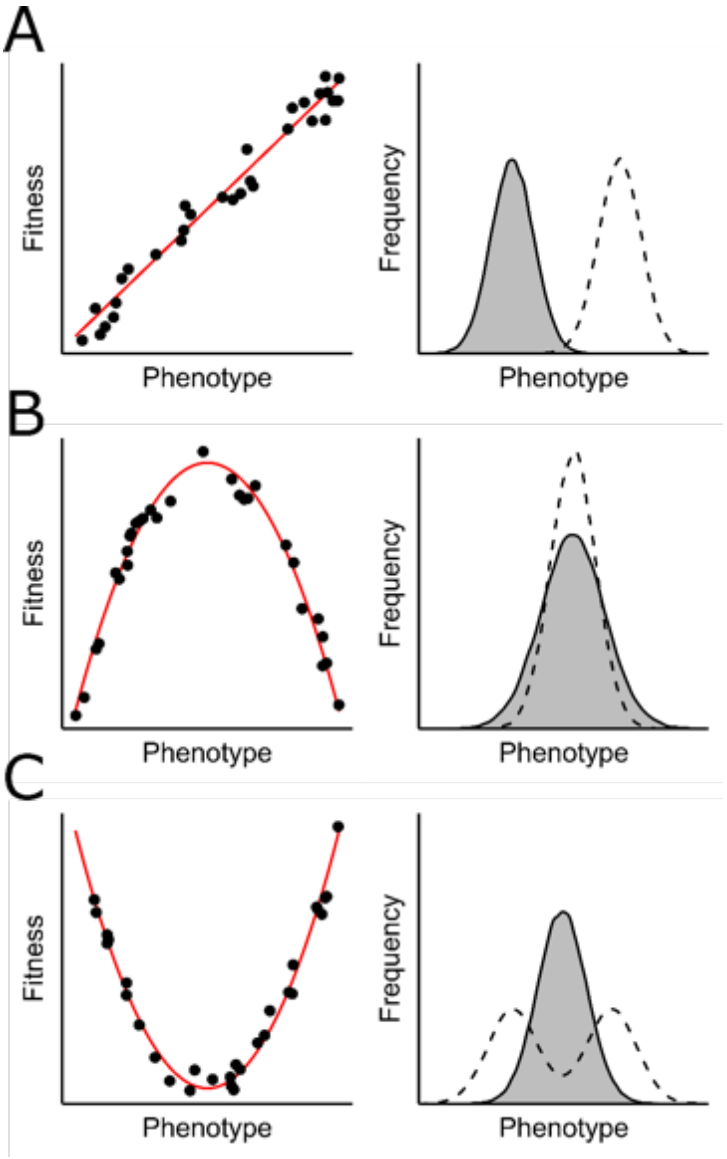


Figure S5.1. Measuring natural selection using selection gradients (Lande and Arnold 1983). Left are hypothetical fitness surfaces for three evolutionary scenarios, and right are histograms of hypothetical changes in population demographics under these respective selection regimes: (A) positive directional selection; (B) stabilising selection; and (C) disruptive selection. Grey is the population before selection, and dashed white lines represent the population after selection.

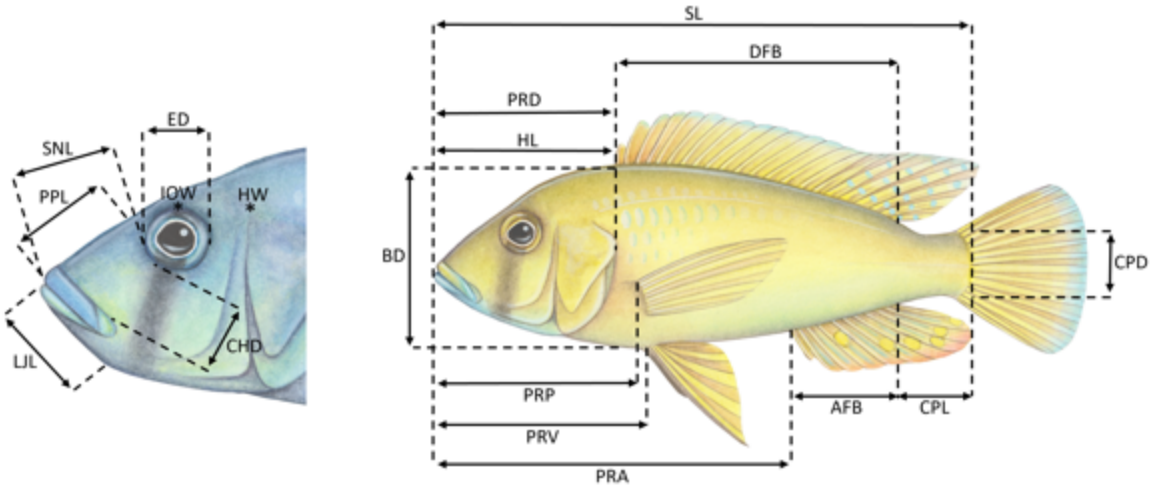


Figure S5.2. The linear morphometric measurements used in this study. SNL = snout length; the distance from the anterior limit of the upper jaw to the bony orbit border. PPL = premaxillary pedicel length; anterior limit of the upper jaw to the tip of the premaxilla's ascending process. LJL = lower jaw length; the full length of the lower jaw. ED = eye diameter; maximum length from the anterior to the posterior tip of the bony orbit. IOW = interorbital width; minimum width of the dossal margin of the orbit. HW = head width; the width of the body at its maximum. CHD = cheek depth; the distance from the most ventral point of the bony orbit to the dorsal-most corner of the supensorium articulation. SL = standard length, the distance from the tip of the snout to the midpoint of the caudal peduncle. DFB = dorsal fin base length, the distance between the most anterior to the most posterior point of the dorsal fin base. PRD = predorsal distance; from the anterior limit of the upper jaw to the anterior limit of the dorsal fin base. HL = head length; distance from the anterior limit of the upper jaw to the most anterior point of the operculum. BD = body depth; the depth of the body at its maximum. PRP = prepectoral distance; from the anterior limit of the upper jaw to the anterior limit of the pectoral fin base. PRV = prepelvic distance; from the anterior limit of the upper jaw to the anterior limit of the pelvic fin base. PRA = preanal distance; from the anterior limit of the upper jaw to the anterior limit of the anal fin base. AFB = anal fin base length; distance from the most anterior to the most posterior point of the anal fin base. CPL = caudal peduncle length; the distance between a vertical line through the most posterior point of the anal fin insertion, through the caudal hypural border. CPD = caudal peduncle depth; the minimum depth of the caudal peduncle.

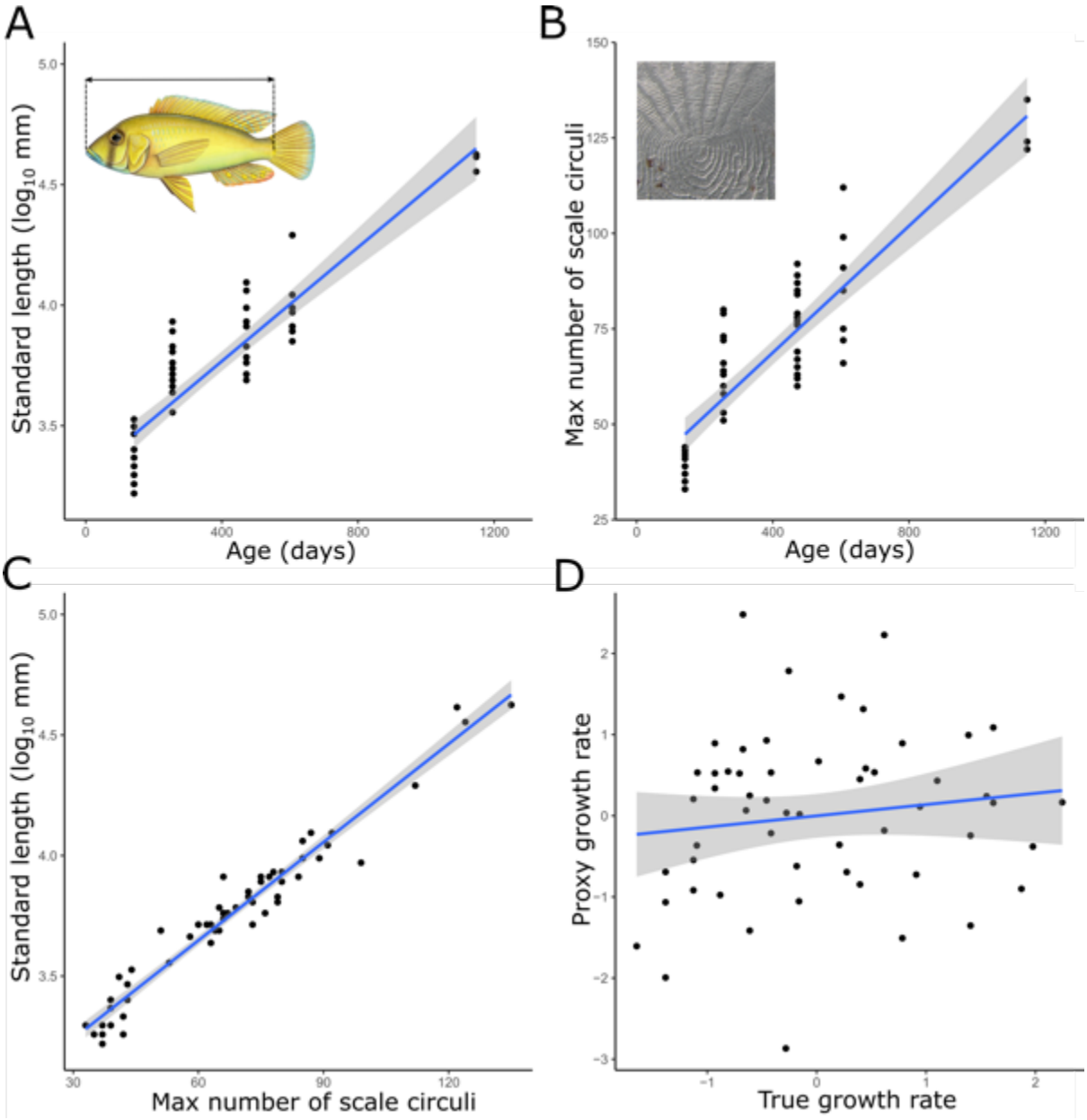


Figure S5.3. Estimating growth rate in *Astatotilapia calliptera*. (A) The relationship between $\log_{10}$ transformed standard length and age. (B) The relationship between the max number of scale circuli and age. (C) The relationship between $\log_{10}$ transformed standard length and the max number of scale circuli. (D) The relationship between our proxy for growth rate, determined by using the max number of scale circuli as a measure of age, and the true growth rate. True growth rate is calculated as the residuals of linear regression of fish age against $\log_{10}$ standard length. Proxy growth rate is calculated as residuals of a linear regression of the max number of scale circuli against $\log_{10}$ standard length.

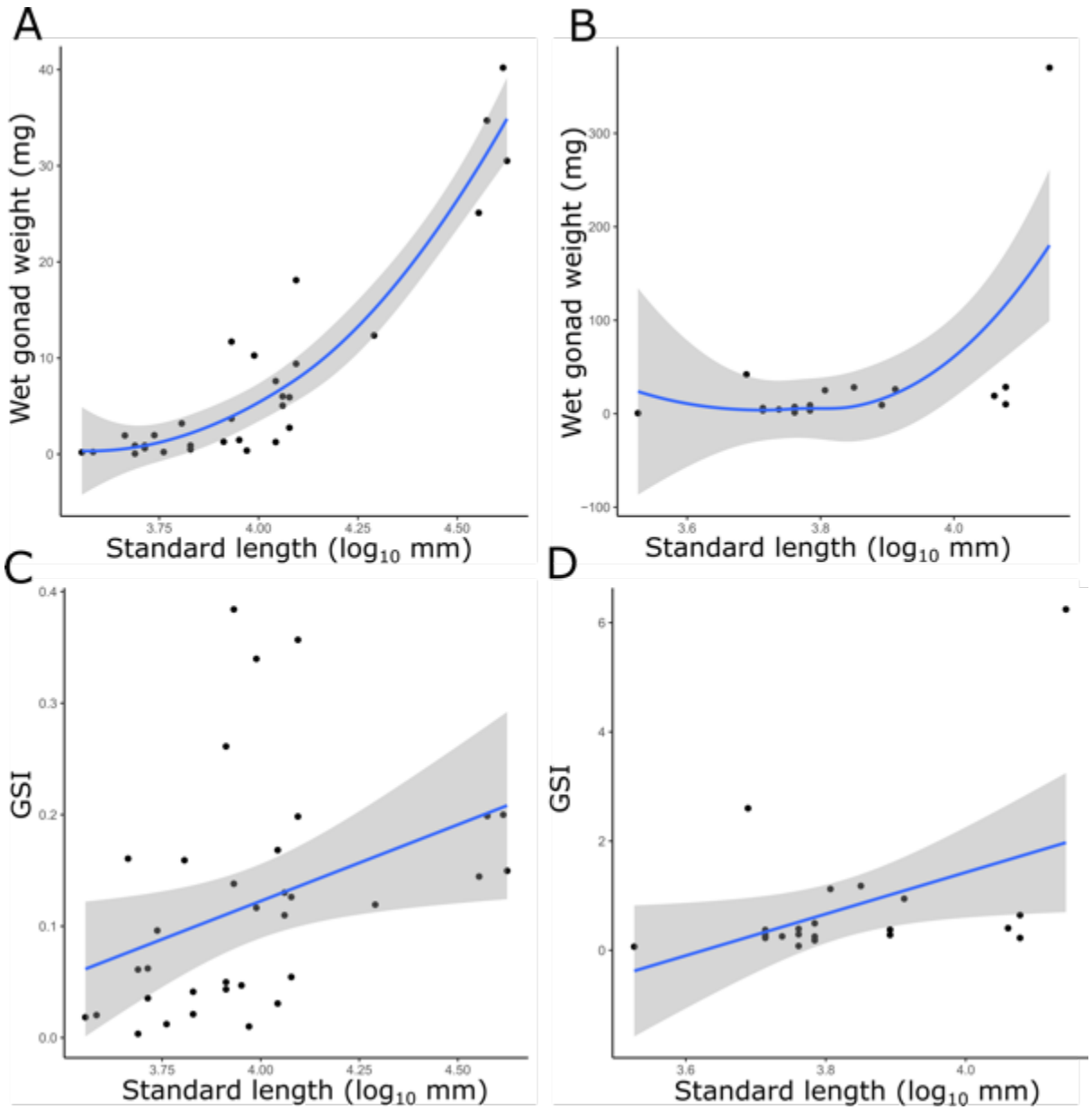


Figure S5.4. The relative gonad weight of *A. calliptera* for males and females. (A) The relationship between $\log_{10}$ standard length and wet gonad weight in male *A. calliptera*. (B) The relationship between $\log_{10}$ standard length and wet gonad weight in female *A. calliptera*. (C) The relationship between log10 standard length and gonadosomatic index (GSI) in male *A. calliptera*. (D) The relationship between $\log_{10}$ standard length and gonadosomatic index (GSI) in female *A. calliptera*. The large female individual with a high gonad weight and GSI was carrying many eggs.

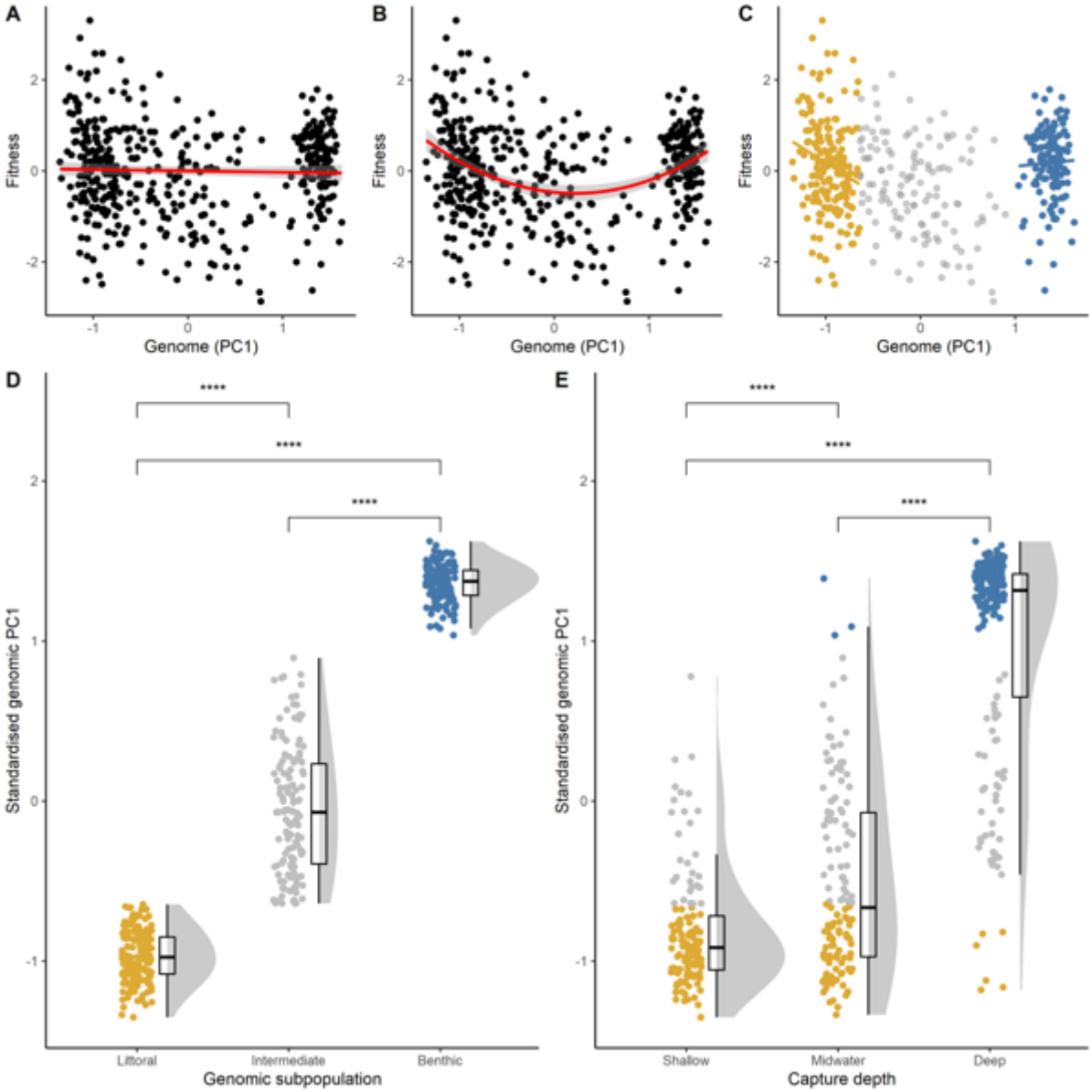


Figure S5.5. Genomic PC1 in Lake Masoko. (A-C) Fitness surfaces for Genomic PC1, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised genomic PC1 values for each (D) genomic subpopulation, and (E) capture depth, are shown. n.b. the subpopulations in D are defined genomically according to genomic PC1 score, so complete lack of overlap between groups here is to be expected. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

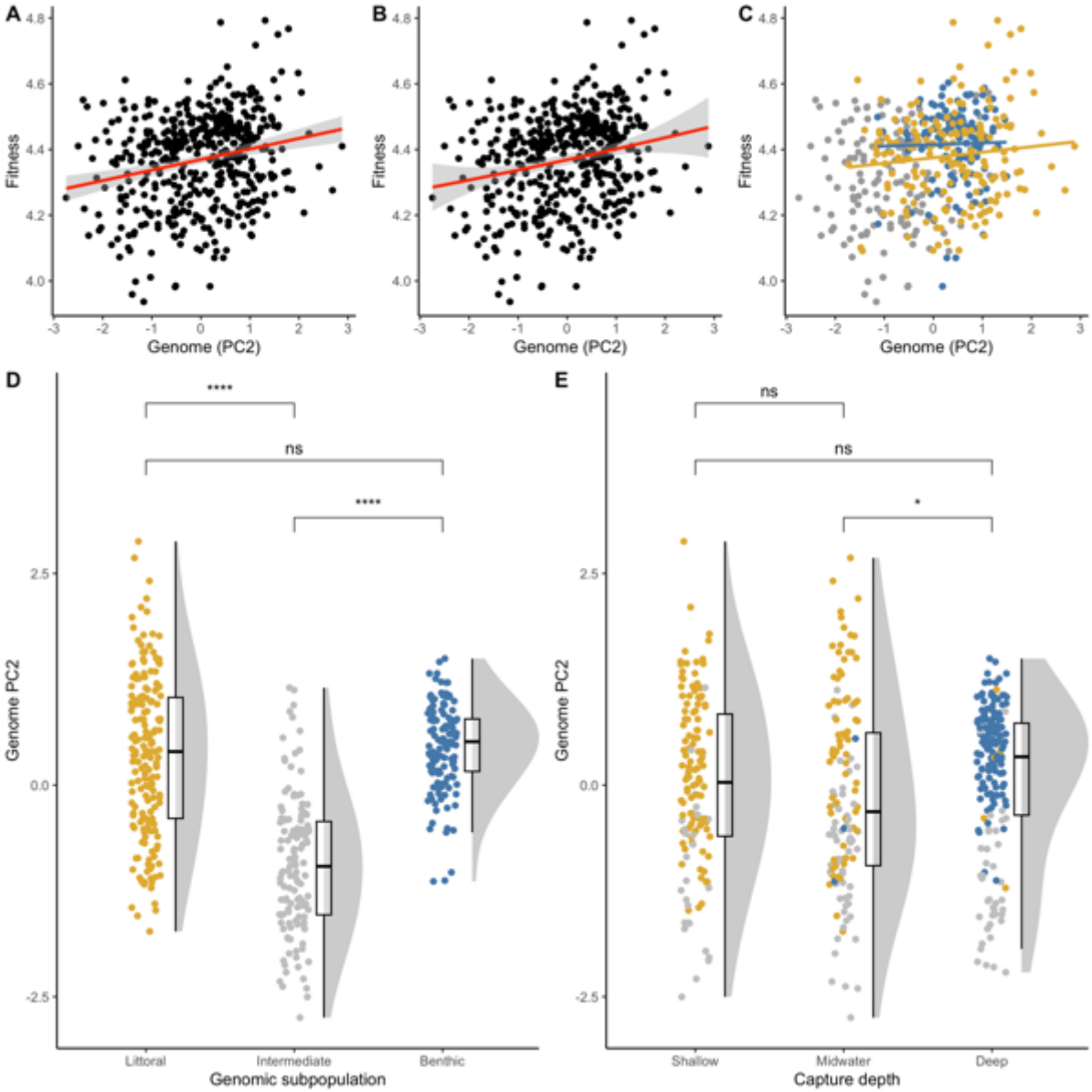


Figure S5.6. Genomic PC2 in Lake Masoko. (A-C) Fitness surfaces for Genomic PC2, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised genomic PC2 values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

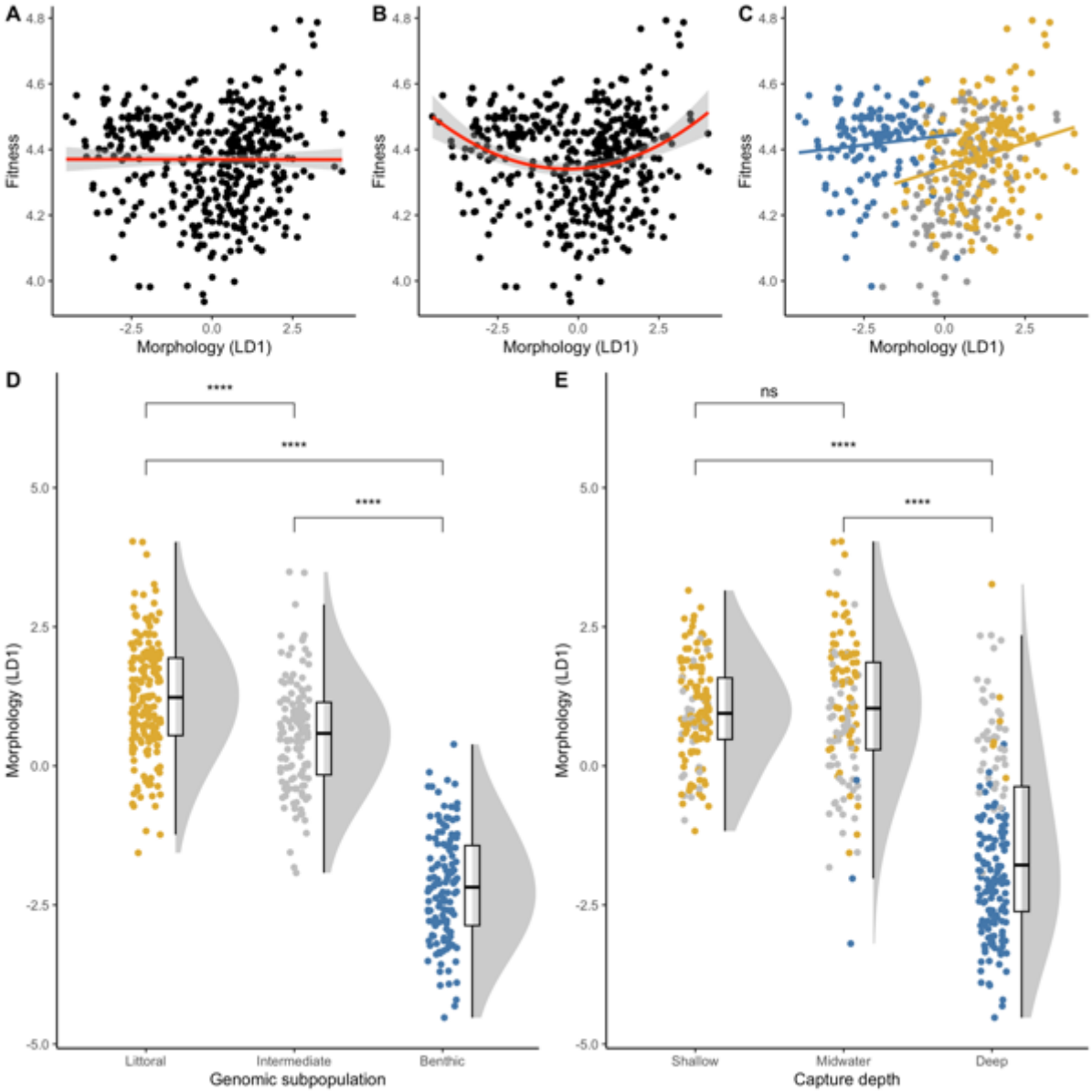


Figure S5.7. Morphological LD1 in Lake Masoko. (A-C) Fitness surfaces for LD1, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) LD1 values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

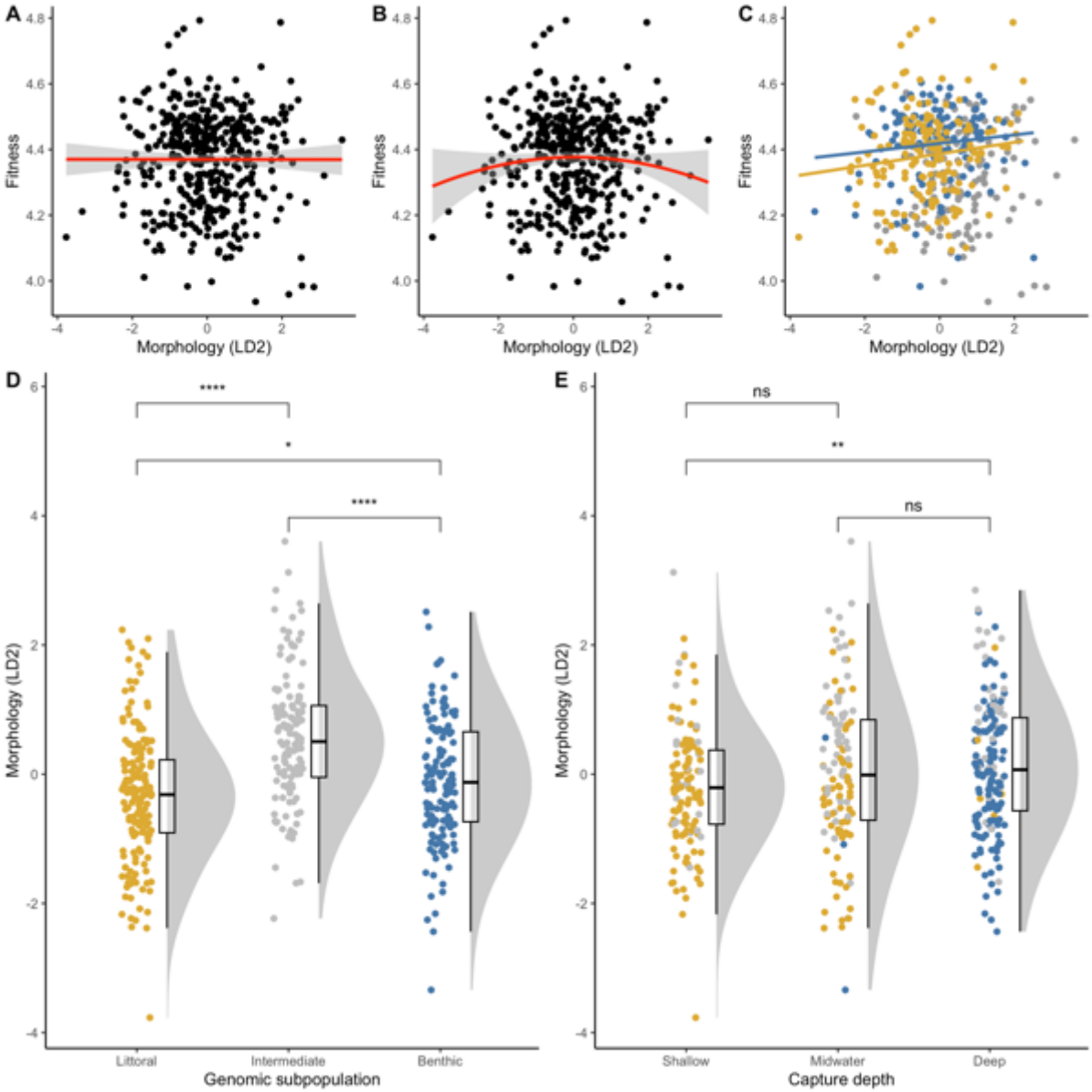


Figure S5.8. Morphological LD2 in Lake Masoko. (A-C) Fitness surfaces for LD2, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) LD2 values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

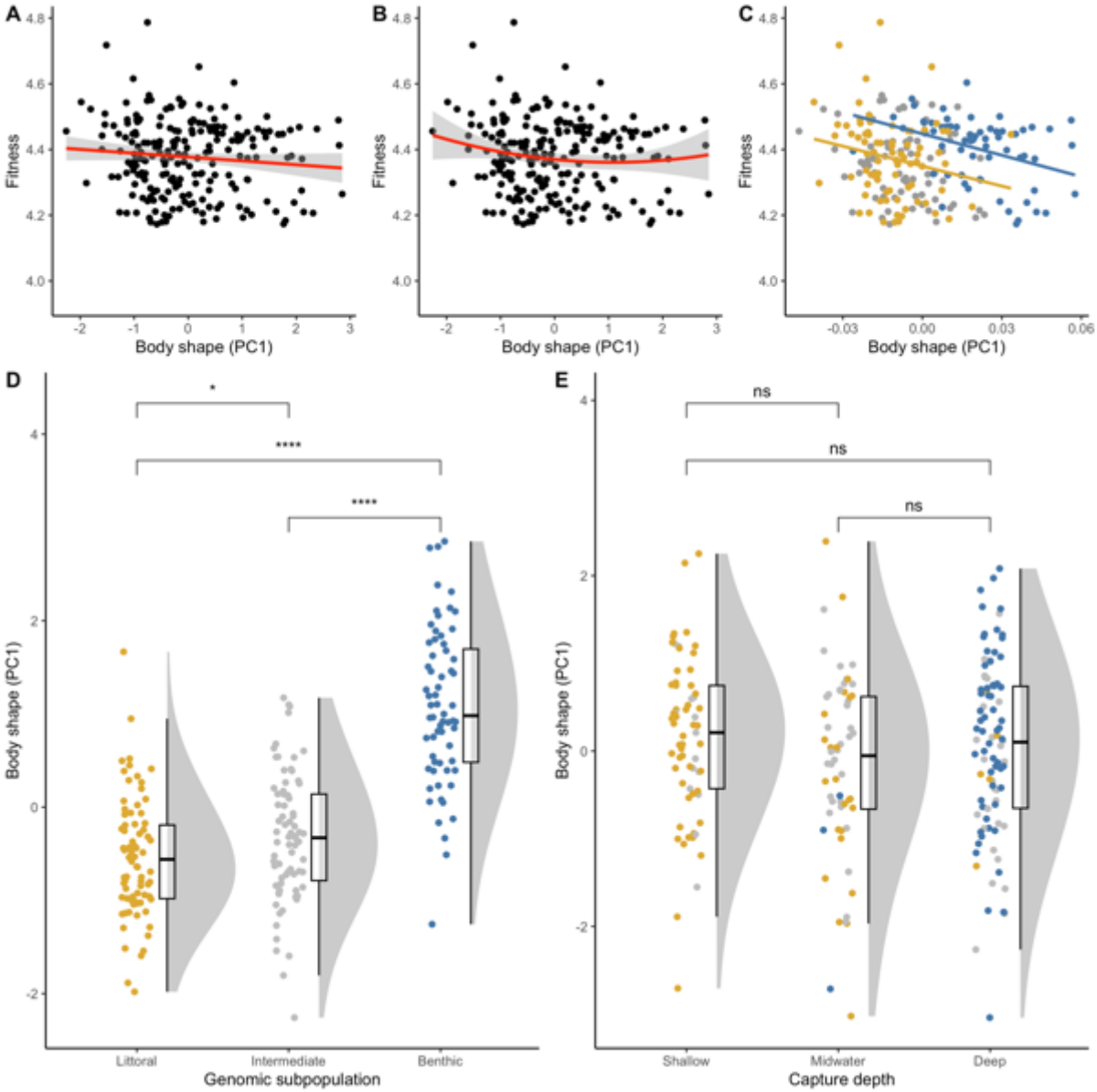


Figure S5.9. PC1 of PCA on body shape in Lake Masoko. (A-C) Fitness surfaces for body shape PC1, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) body shape PC1 values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

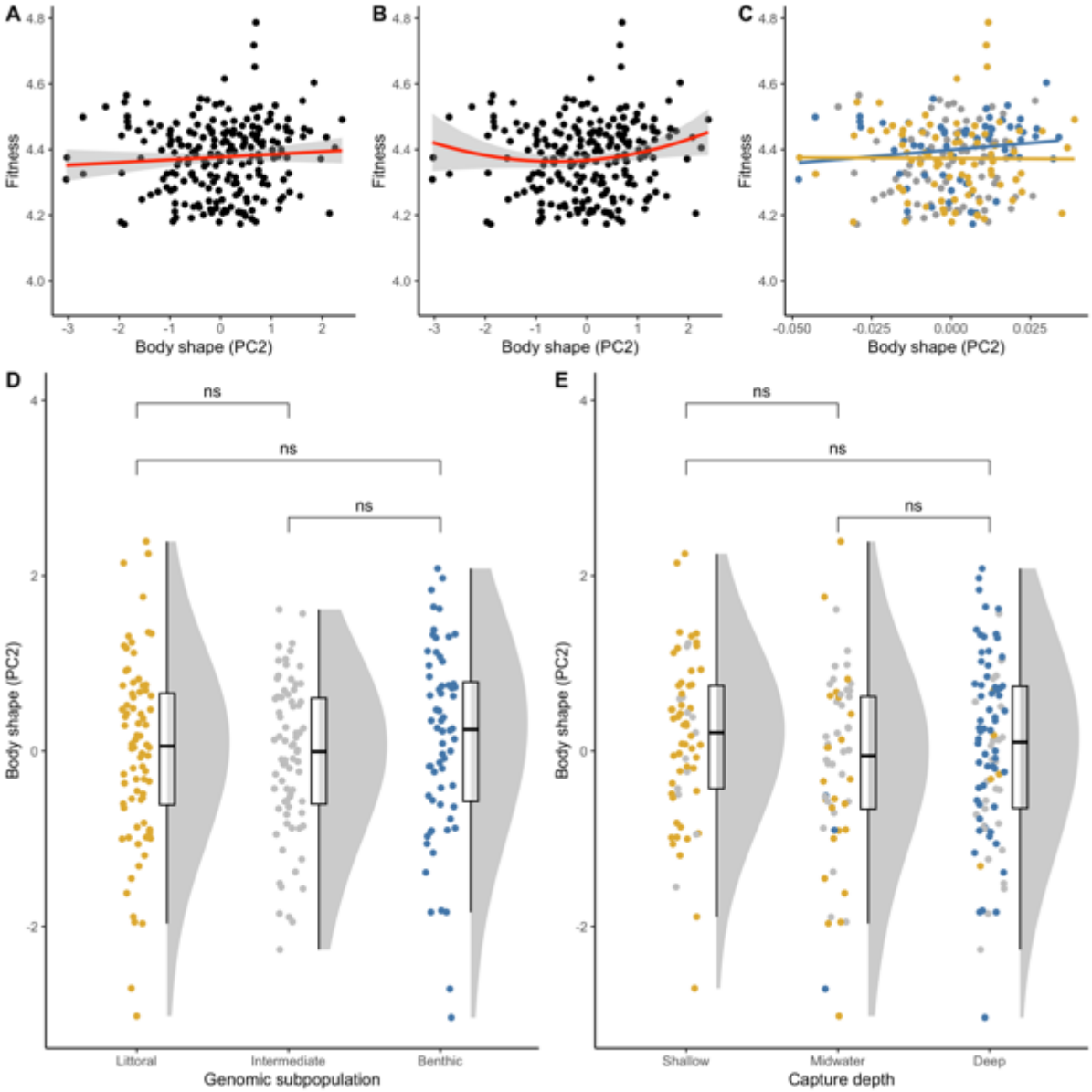


Figure S5.10. PC2 of PCA on body shape in Lake Masoko. (A-C) Fitness surfaces for body shape PC2, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) body shape PC2 values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

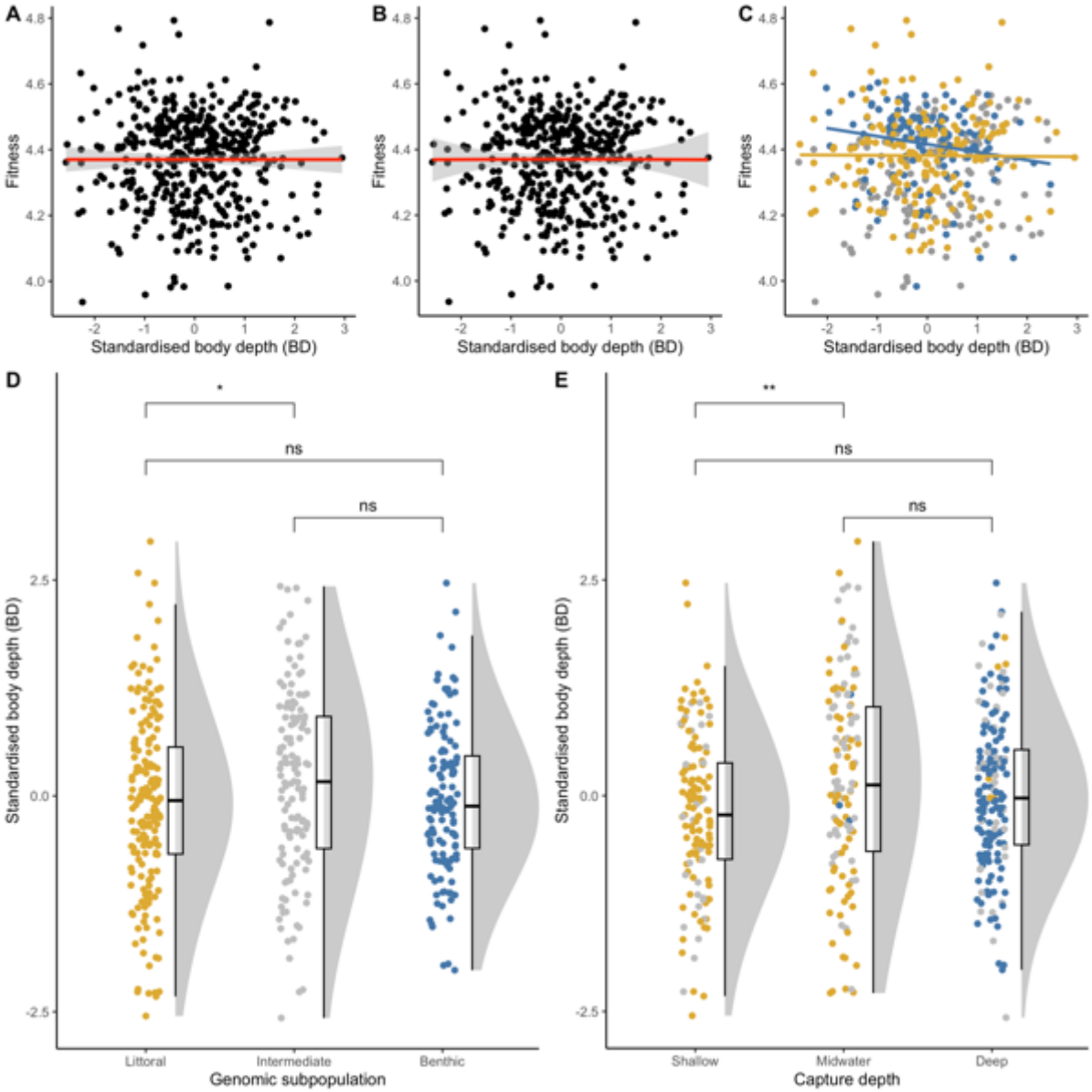


Figure S5.11. Standardised body depth in Lake Masoko. (A-C) Fitness surfaces for standardised body depth, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised body depth values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

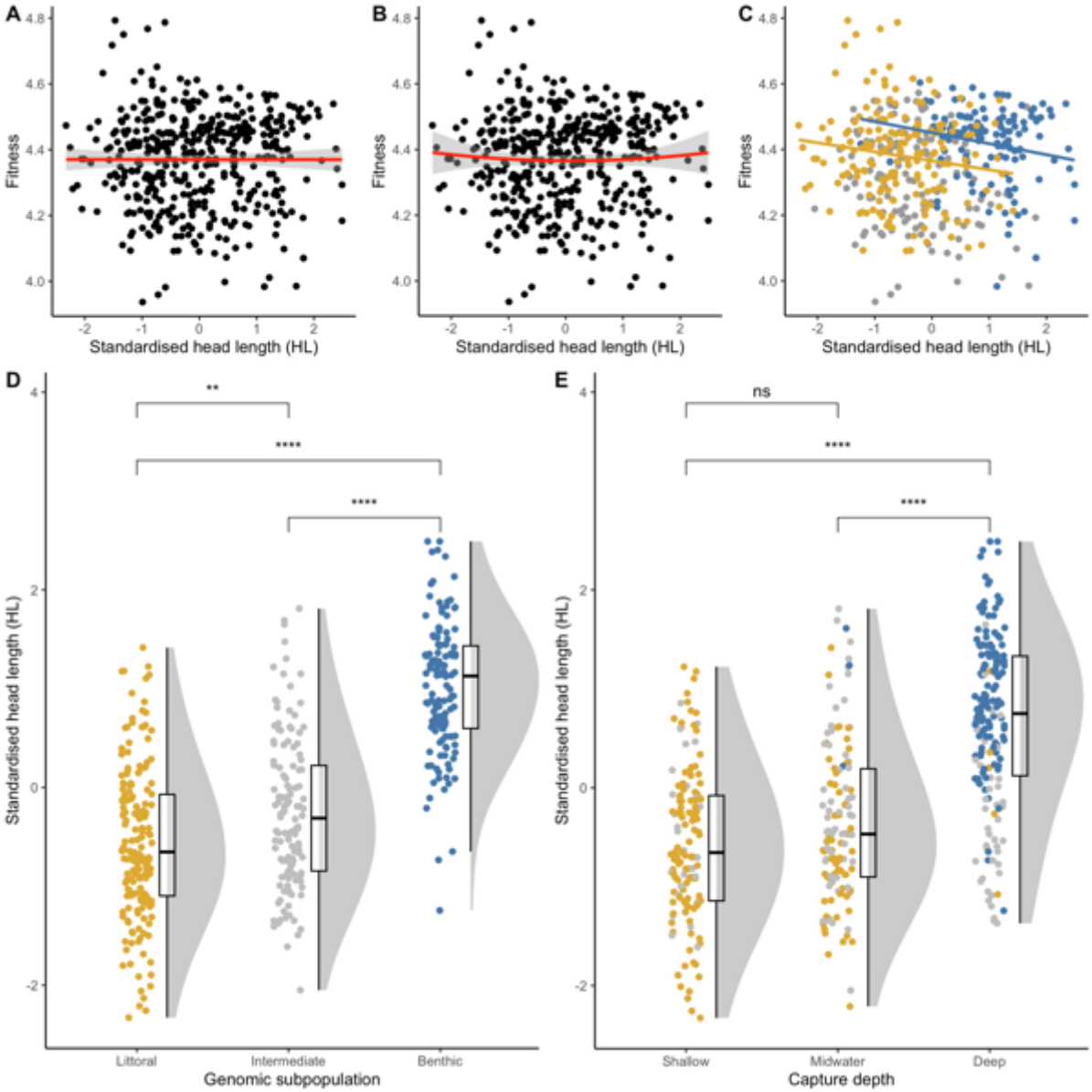


Figure S5.12. Standardised head length in Lake Masoko. (A-C) Fitness surfaces for standardised head length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised head length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

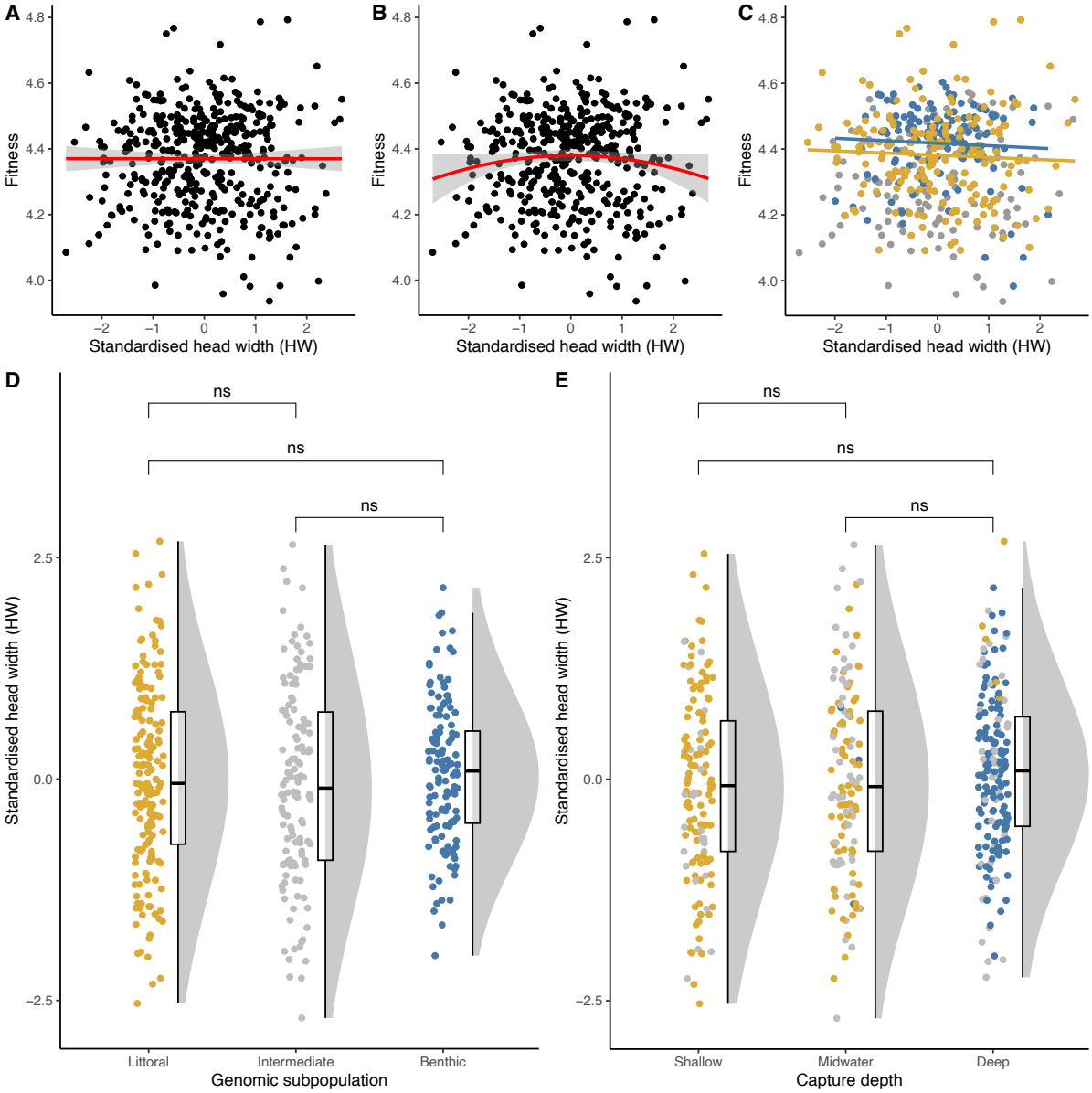


Figure S5.13. Standardised head width in Lake Masoko. (A-C) Fitness surfaces for standardised head width, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised head width values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

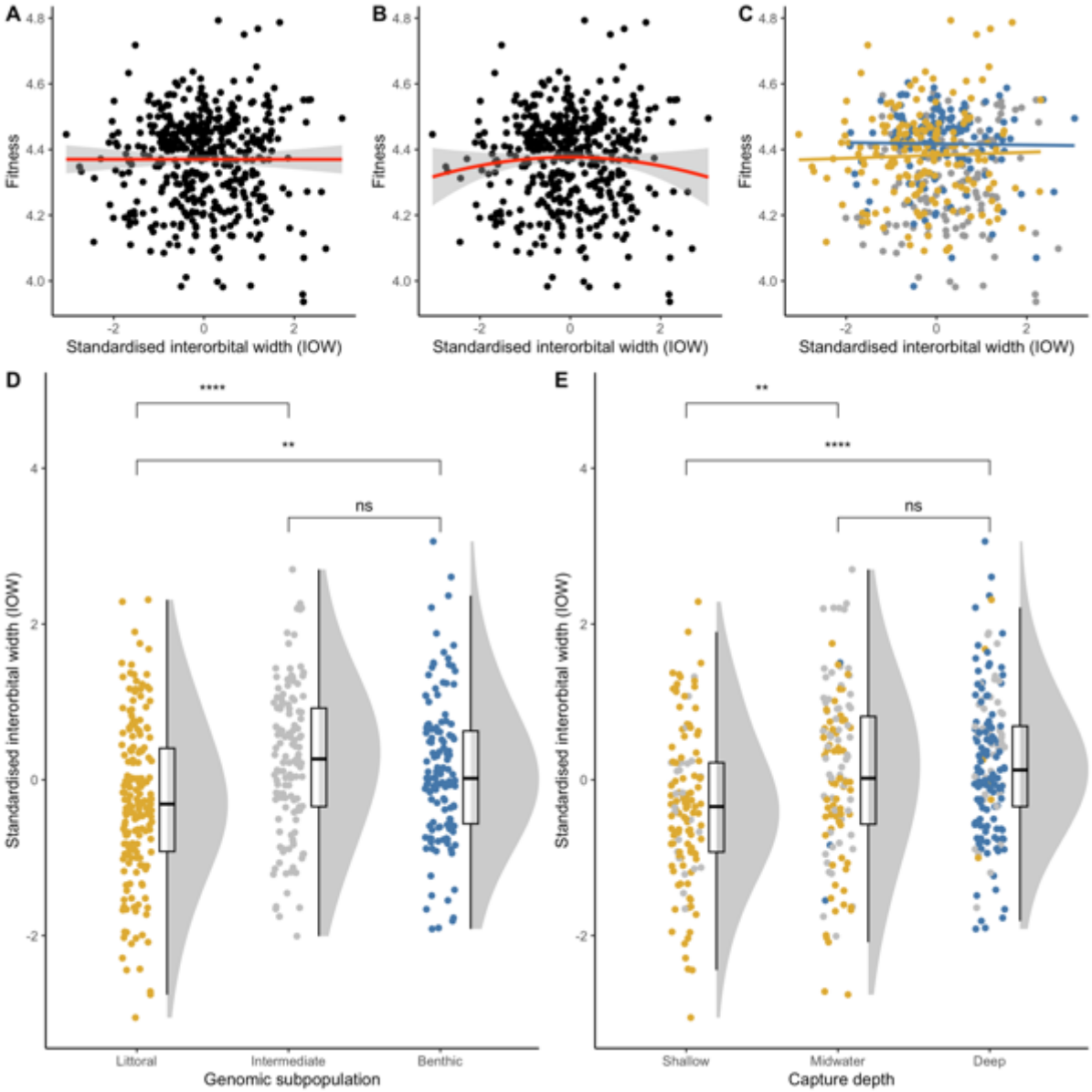


Figure S5.14. Standardised interorbital width in Lake Masoko. (A-C) Fitness surfaces for standardised interorbital width, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised interorbital width values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

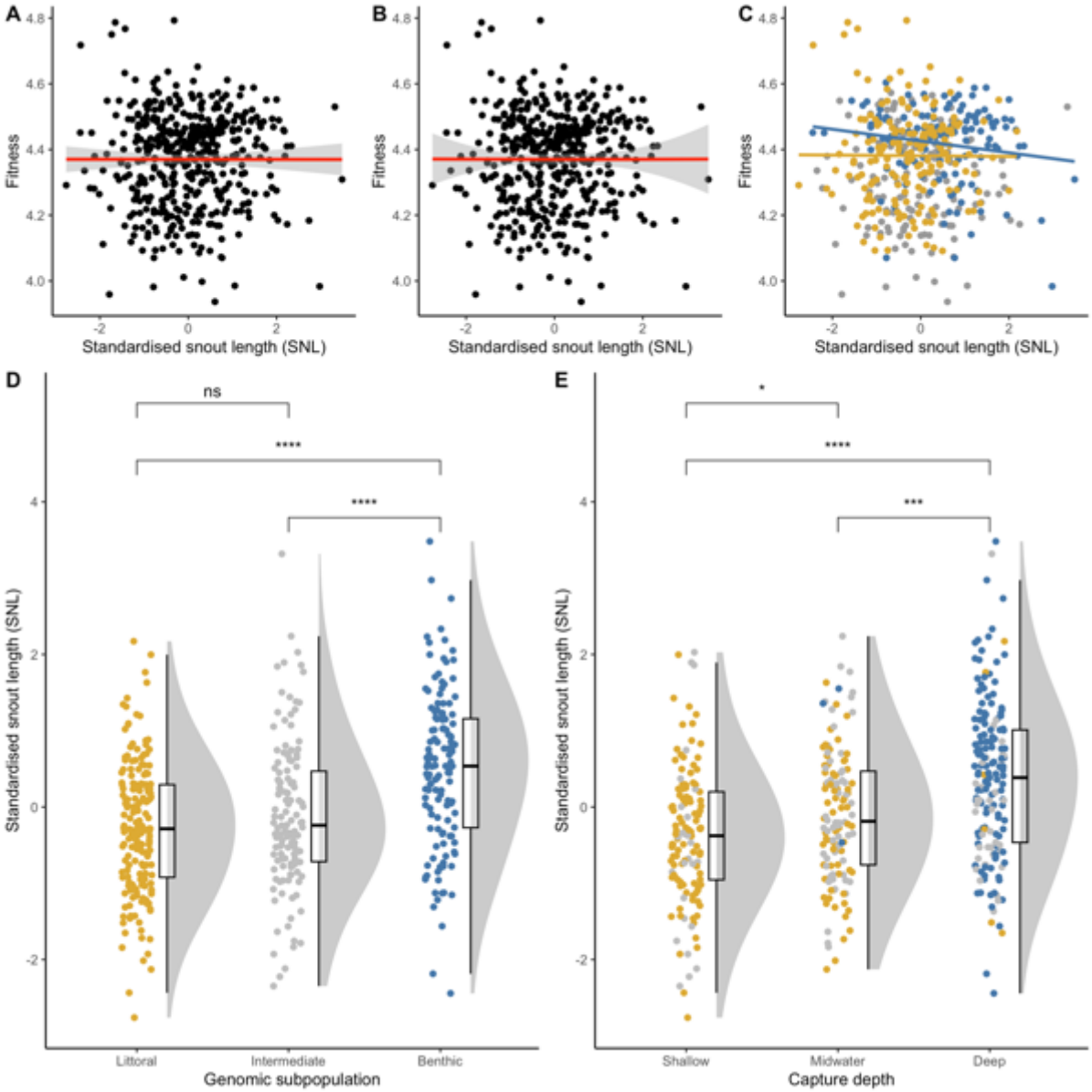


Figure S5.15. Standardised snout length in Lake Masoko. (A-C) Fitness surfaces for standardised snout length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised snout length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

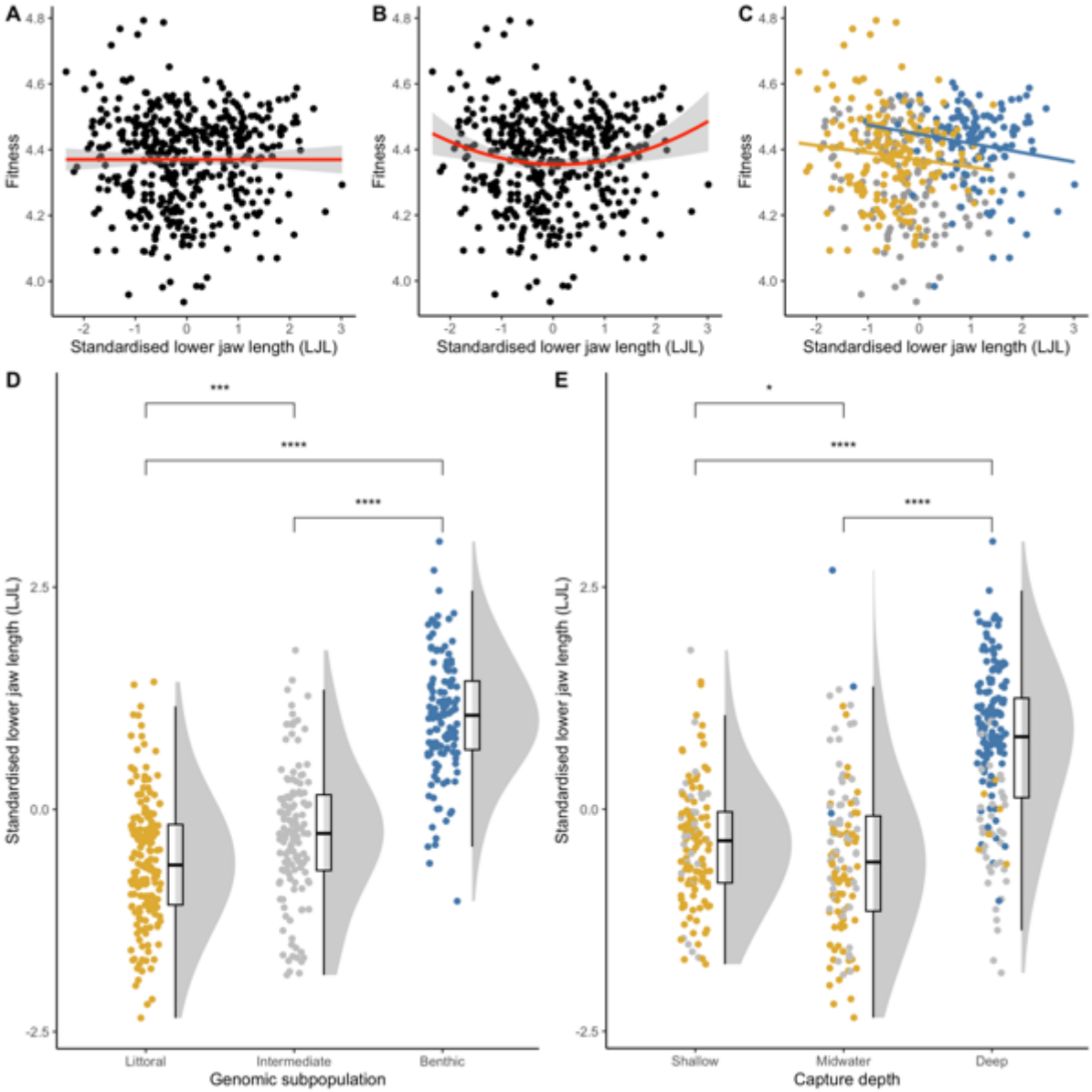


Figure S5.16. Standardised lower jaw length in Lake Masoko. (A-C) Fitness surfaces for standardised lower jaw length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised lower jaw length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

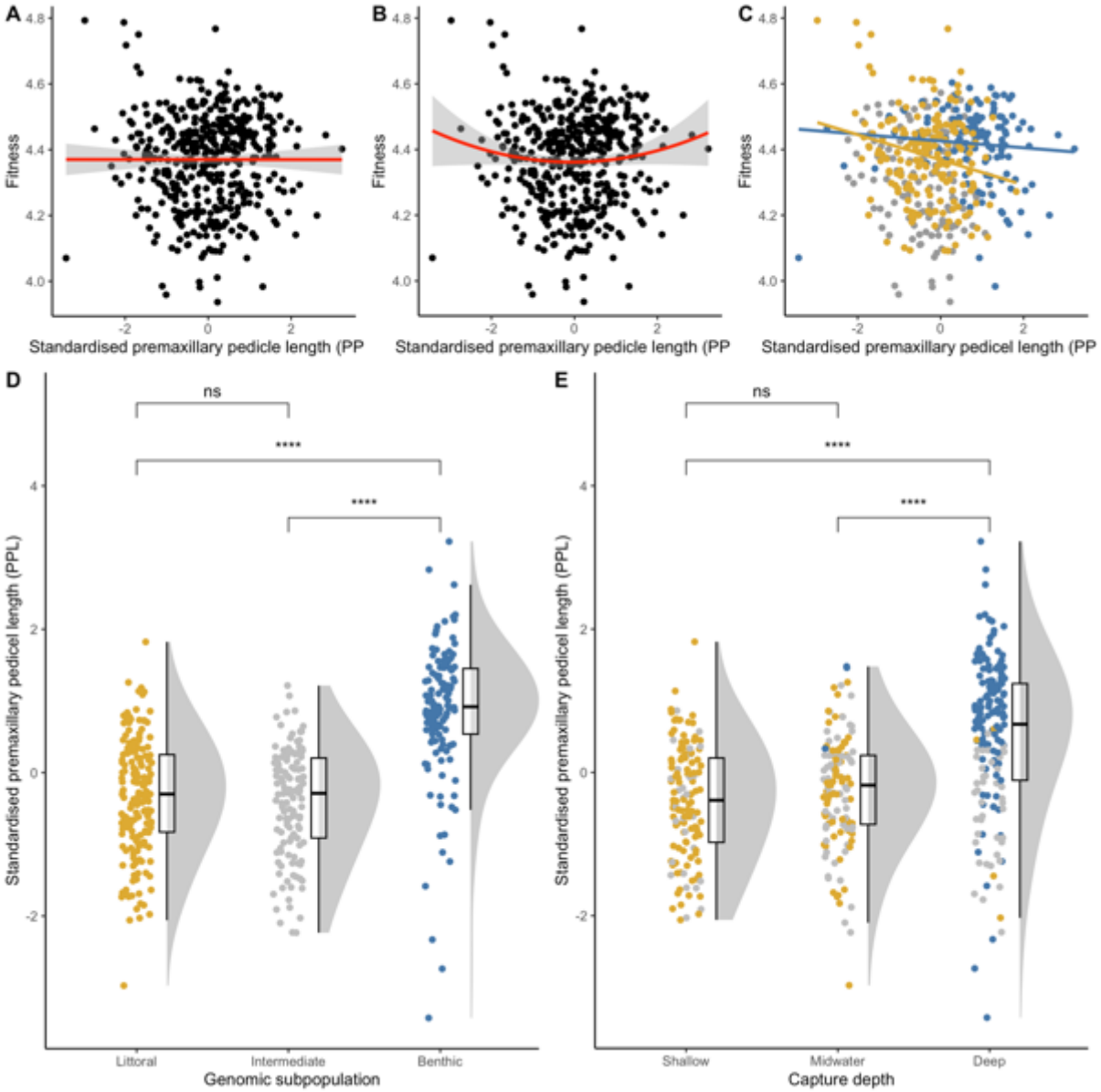


Figure S5.17. Standardised premaxillary pedicel length in Lake Masoko. (A-C) Fitness surfaces for standardised premaxillary pedicel length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised premaxillary pedicel length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

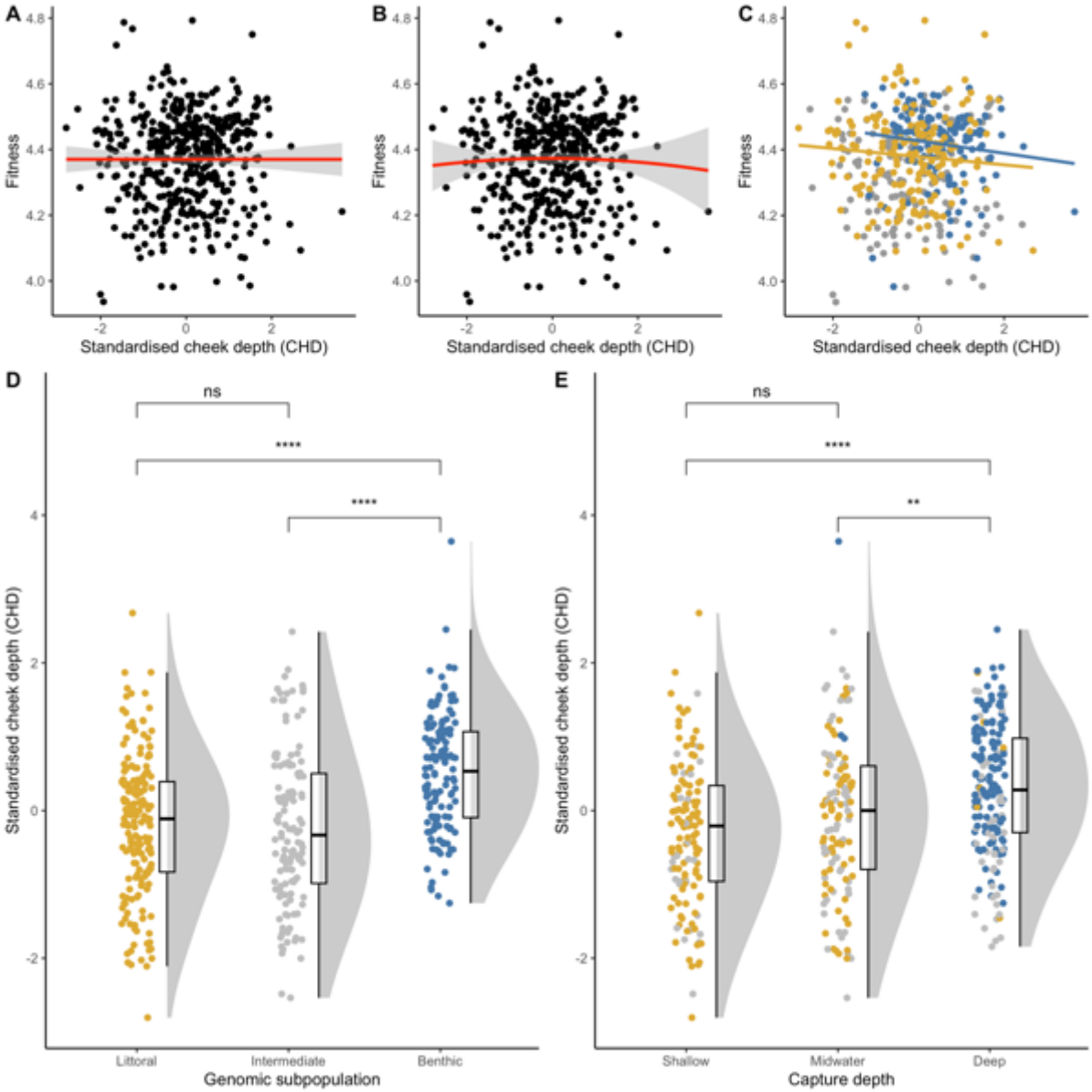


Figure S5.18. Standardised cheek depth in Lake Masoko. (A-C) Fitness surfaces for standardised cheek depth, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised cheek depth values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

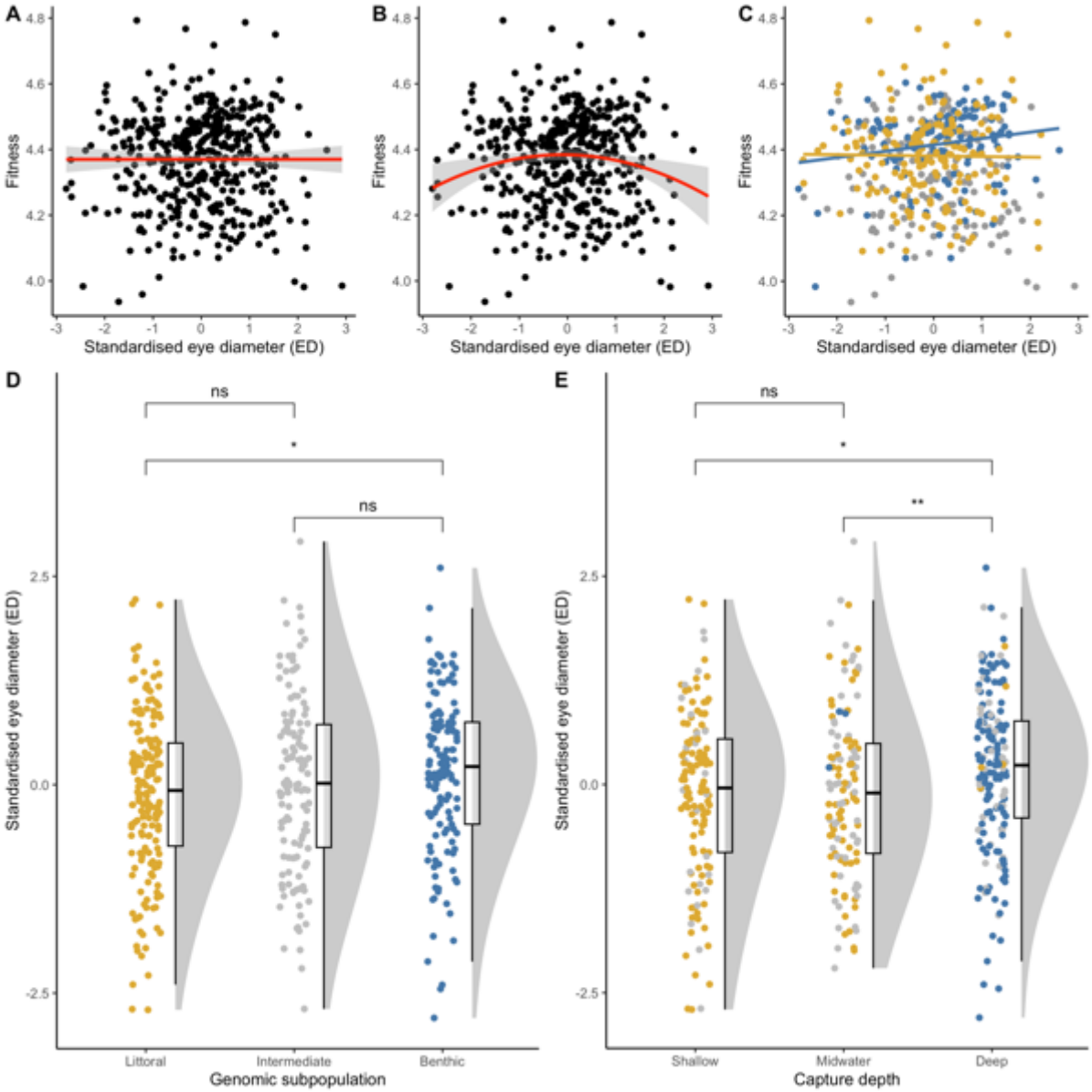


Figure S5.19. Standardised eye diameter in Lake Masoko. (A-C) Fitness surfaces for standardised eye diameter, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised eye diameter values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

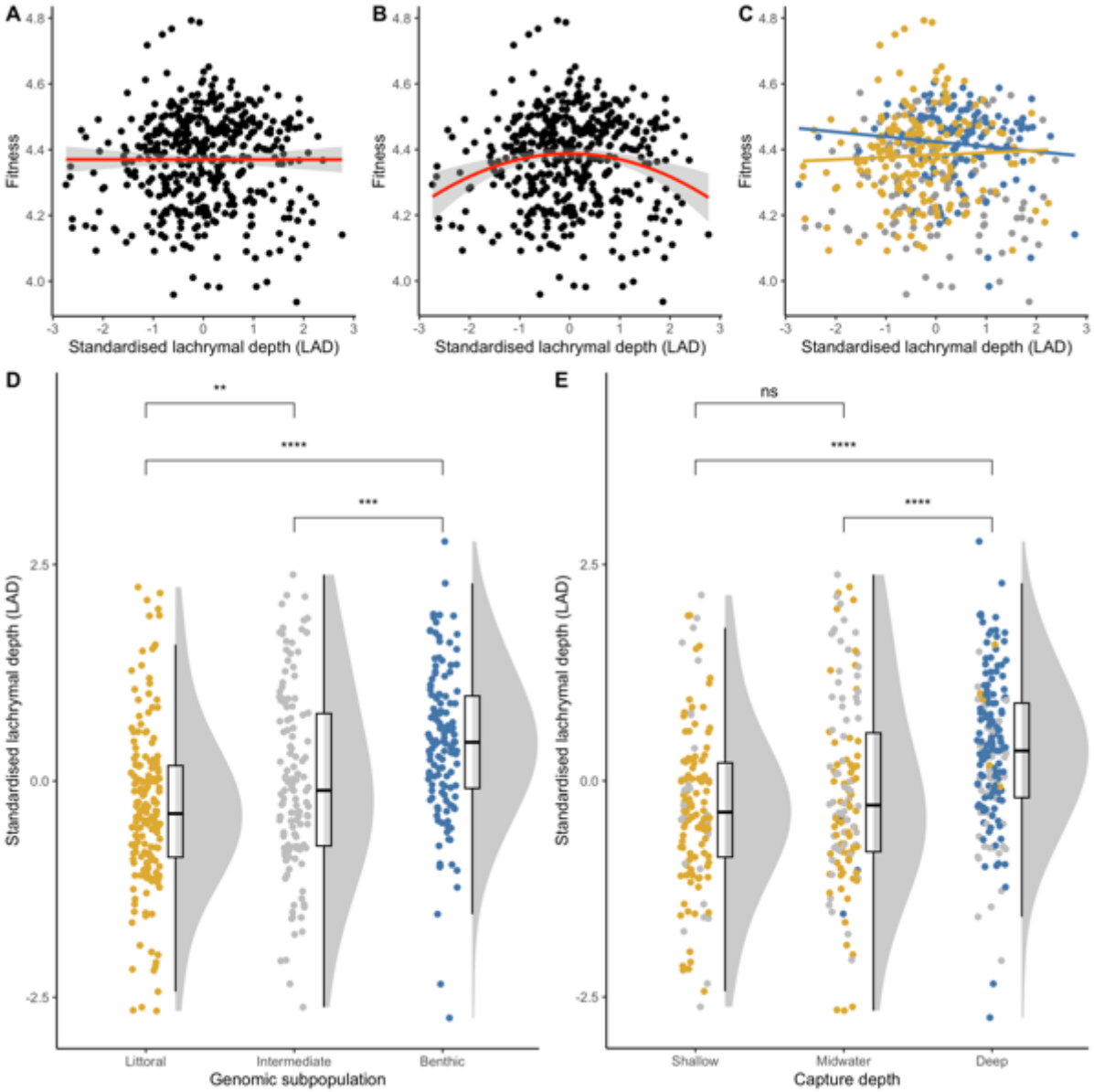


Figure S5.20. Standardised lachrymal depth in Lake Masoko. (A-C) Fitness surfaces for standardised lachrymal depth, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised lachrymal depth values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

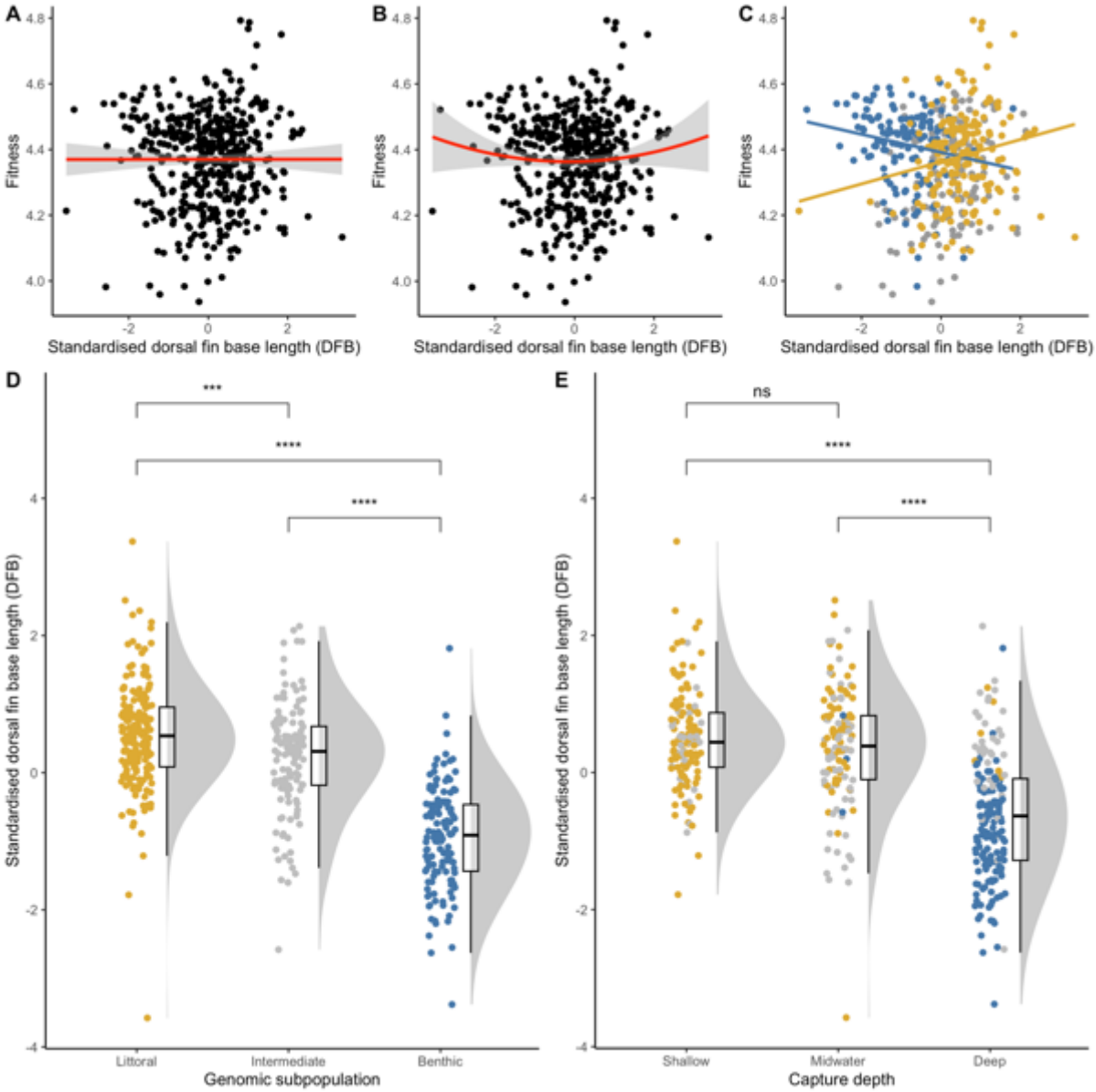


Figure S5.21. Standardised dorsal fin base length in Lake Masoko. (A-C) Fitness surfaces for standardised dorsal fin base length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised dorsal fin base length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

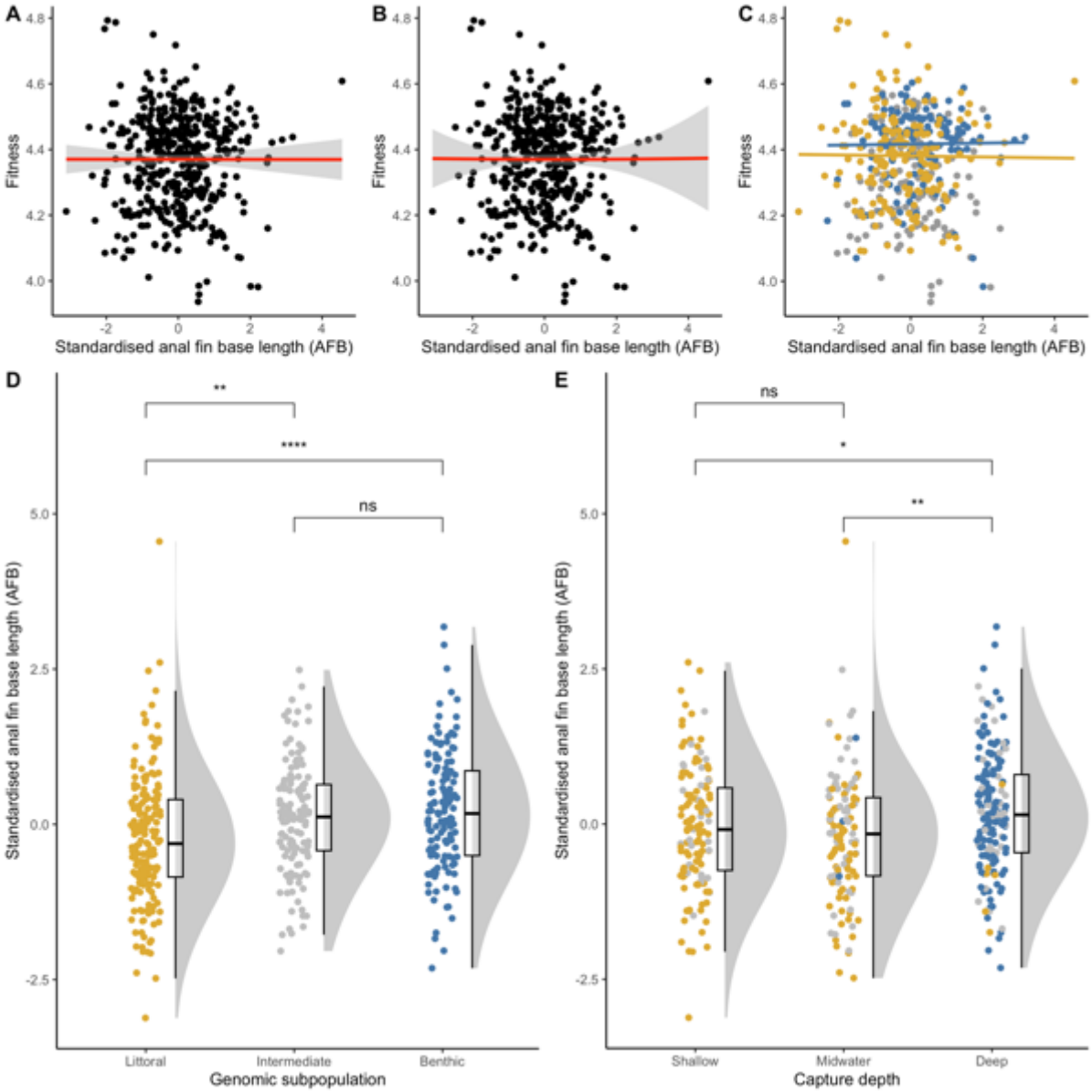


Figure S5.22. Standardised anal fin base length in Lake Masoko. (A-C) Fitness surfaces for standardised anal fin base length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised anal fin base length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

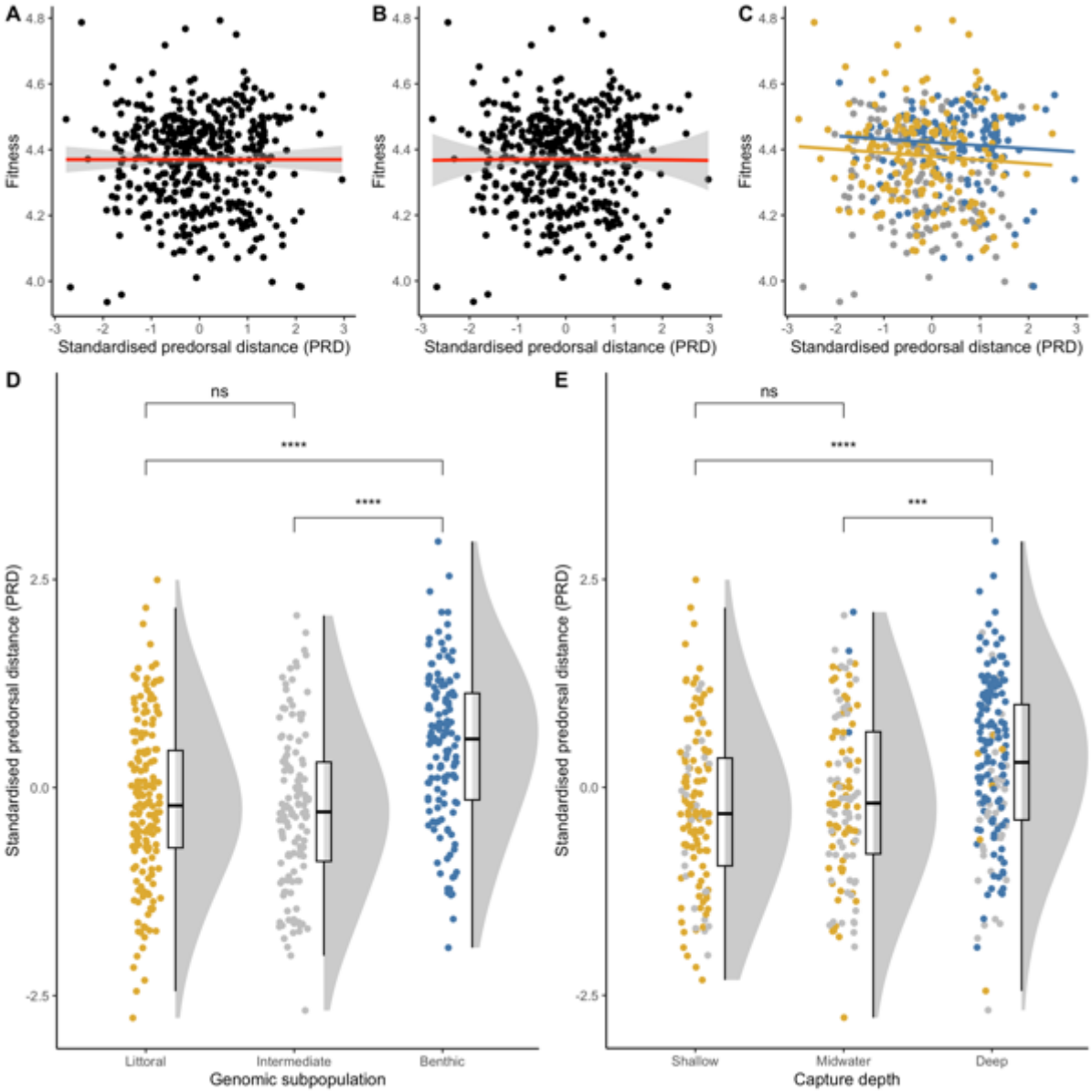


Figure S5.22. Standardised predorsal distance in Lake Masoko. (A-C) Fitness surfaces for standardised predorsal distance, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised predorsal distance values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

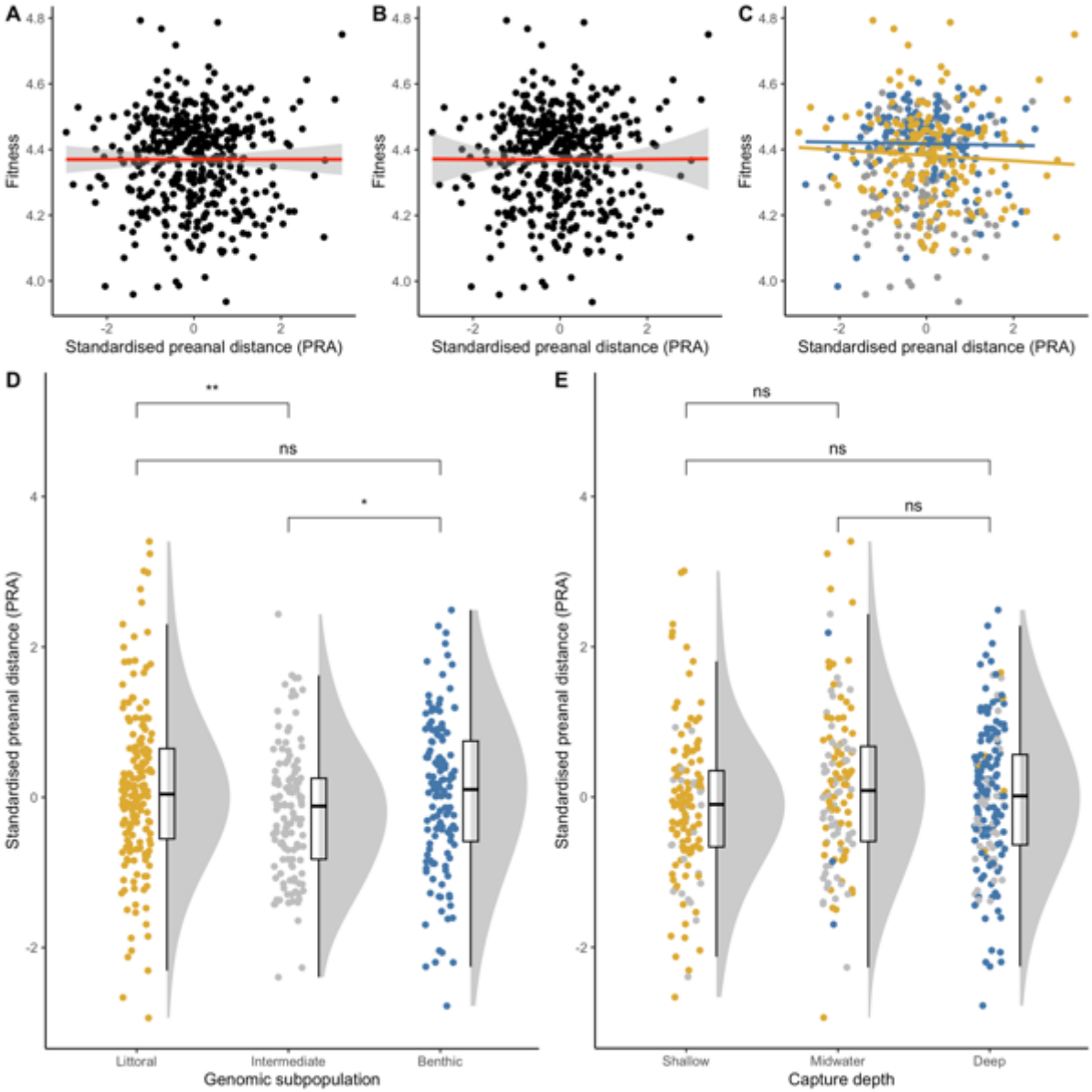


Figure S5.23. Standardised preanal distance in Lake Masoko. (A-C) Fitness surfaces for standardised preanal distance, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised preanal distance values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

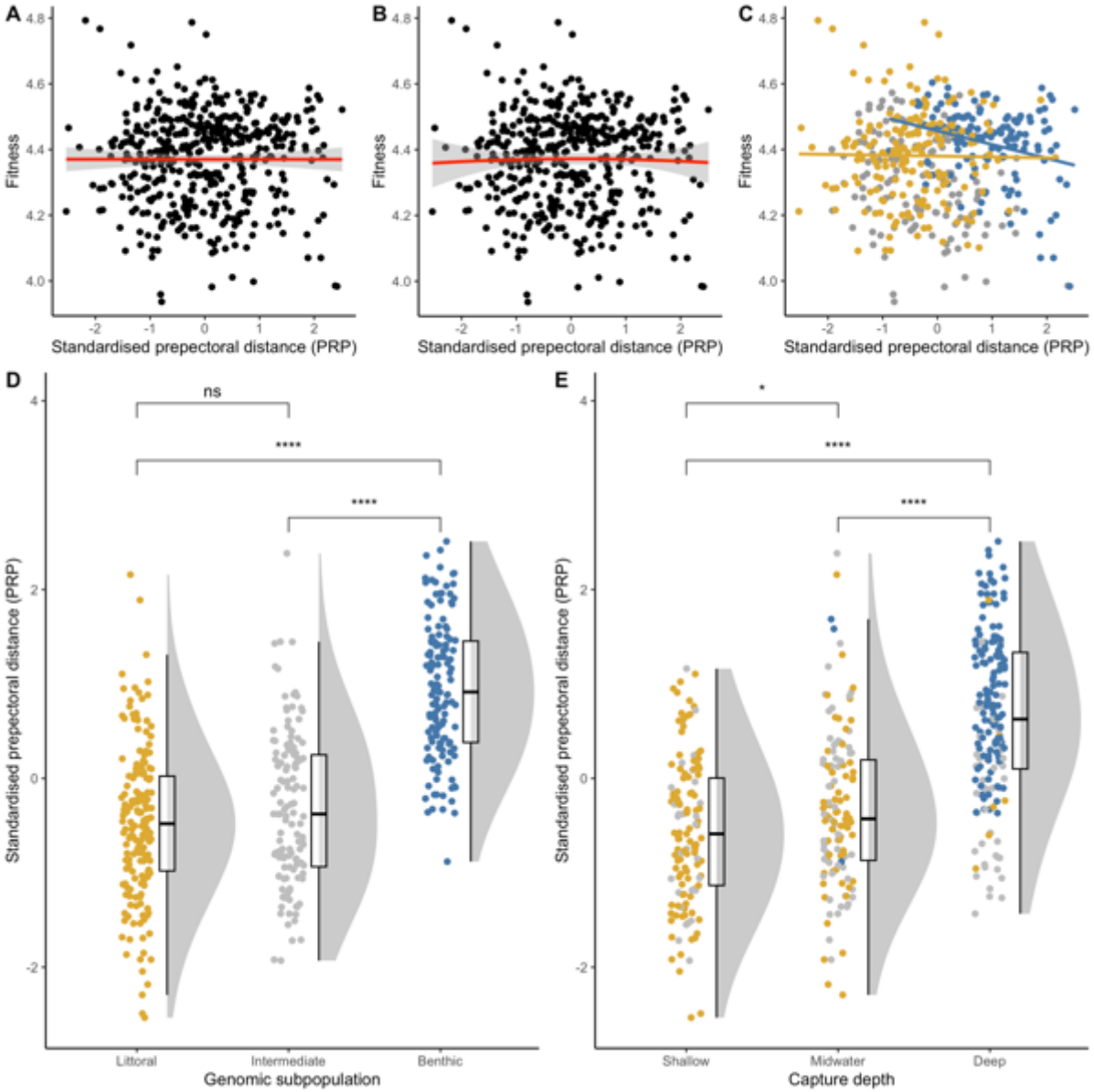


Figure S5.24. Standardised prepectoral distance in Lake Masoko. (A-C) Fitness surfaces for standardised prepectoral distance, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised prepectoral distance values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

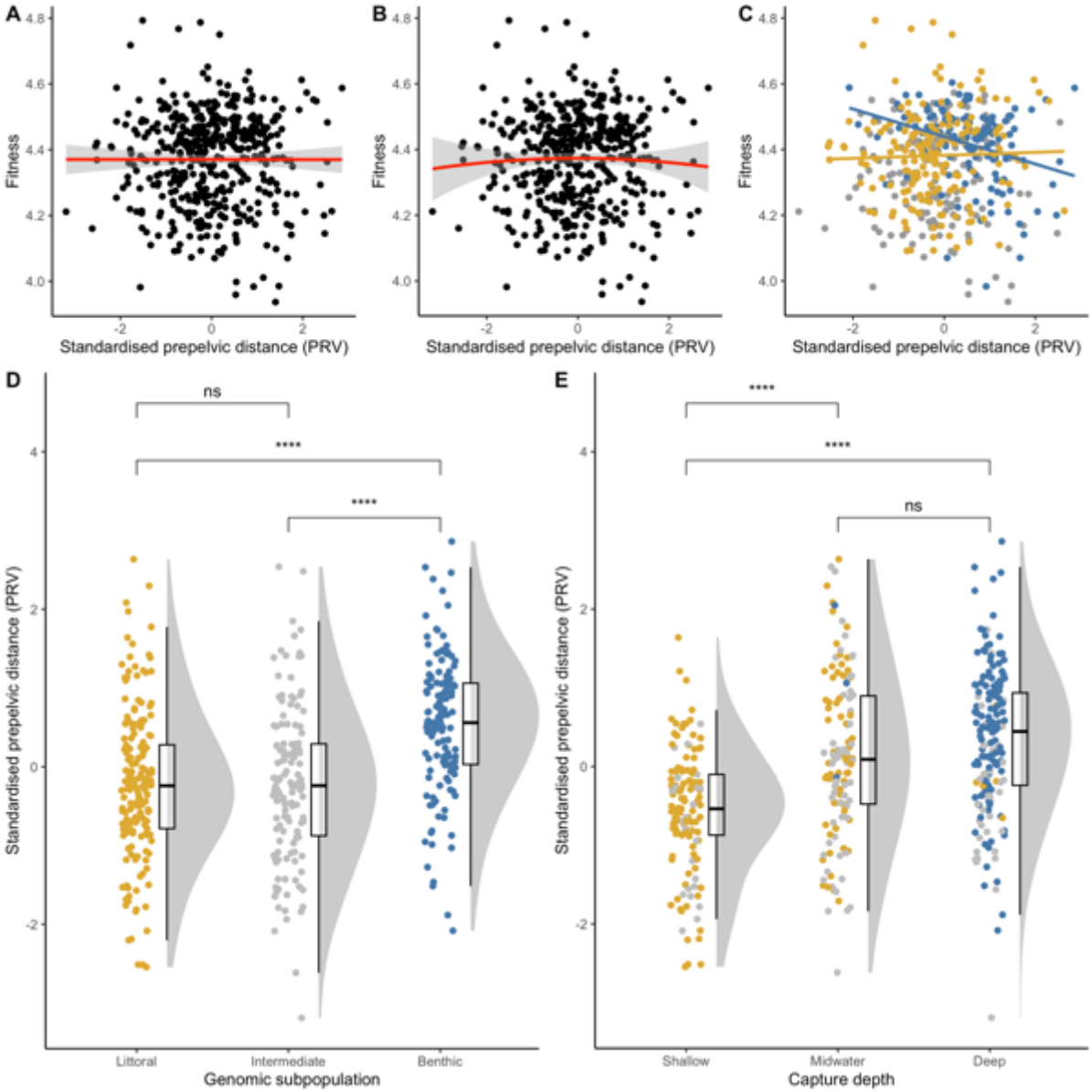


Figure S5.25. Standardised prepelvic distance in Lake Masoko. (A-C) Fitness surfaces for standardised prepelvic distance, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised prepelvic distance values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

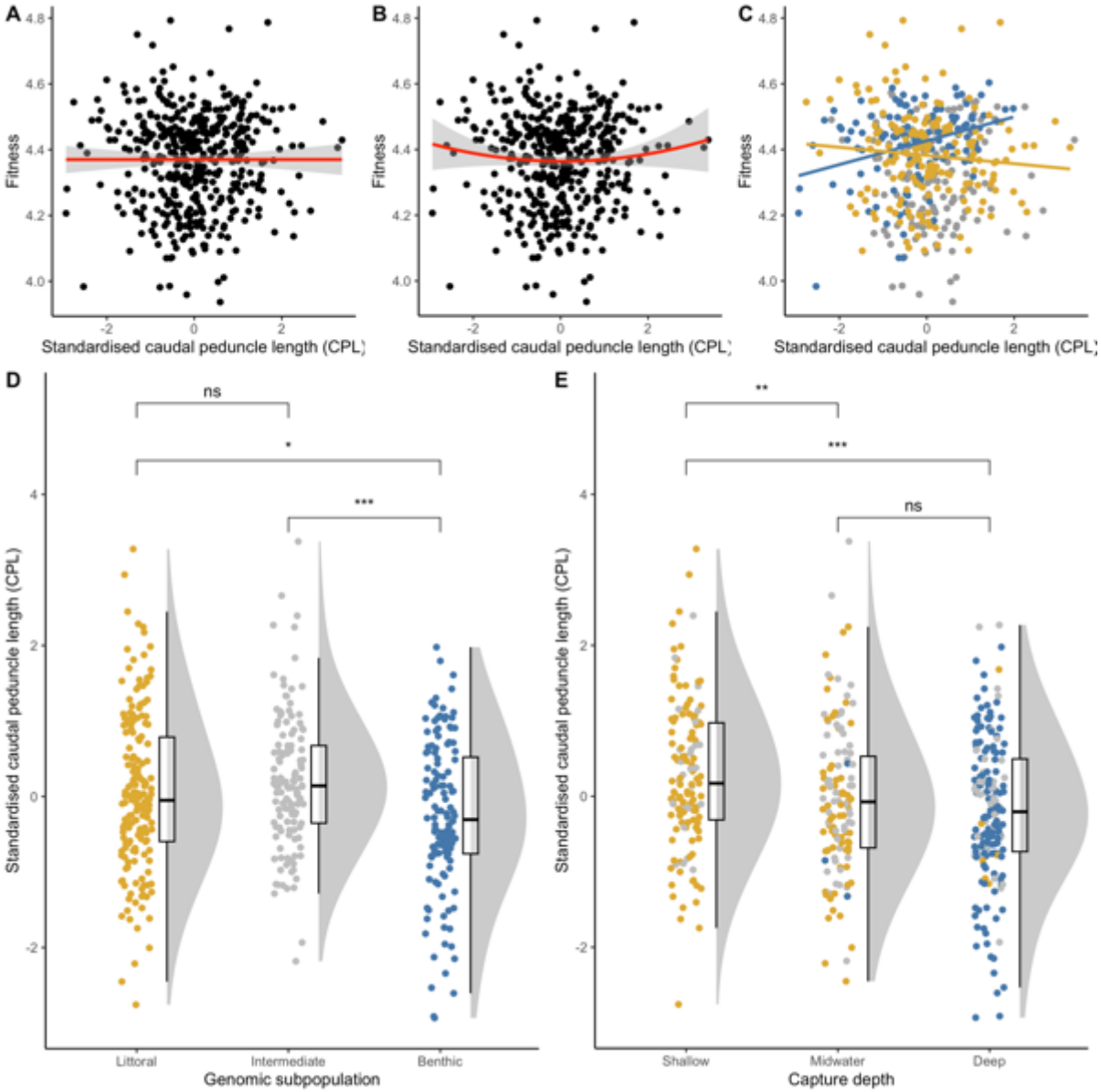


Figure S5.26. Standardised caudal peduncle length in Lake Masoko. (A-C) Fitness surfaces for standardised caudal peduncle length, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised caudal peduncle length values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

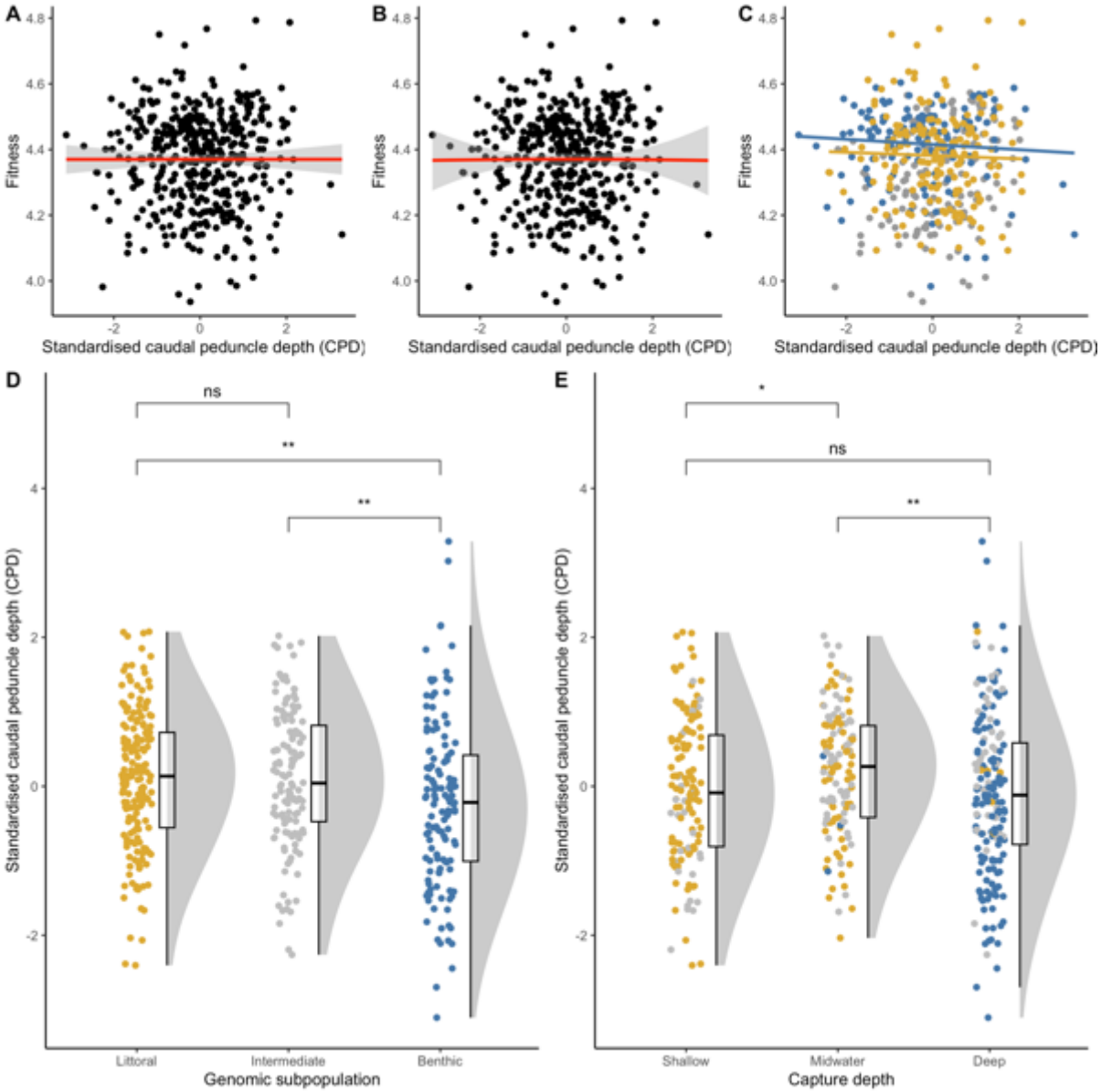


Figure S5.27. Standardised caudal peduncle depth in Lake Masoko. (A-C) Fitness surfaces for standardised caudal peduncle depth, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised caudal peduncle depth values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

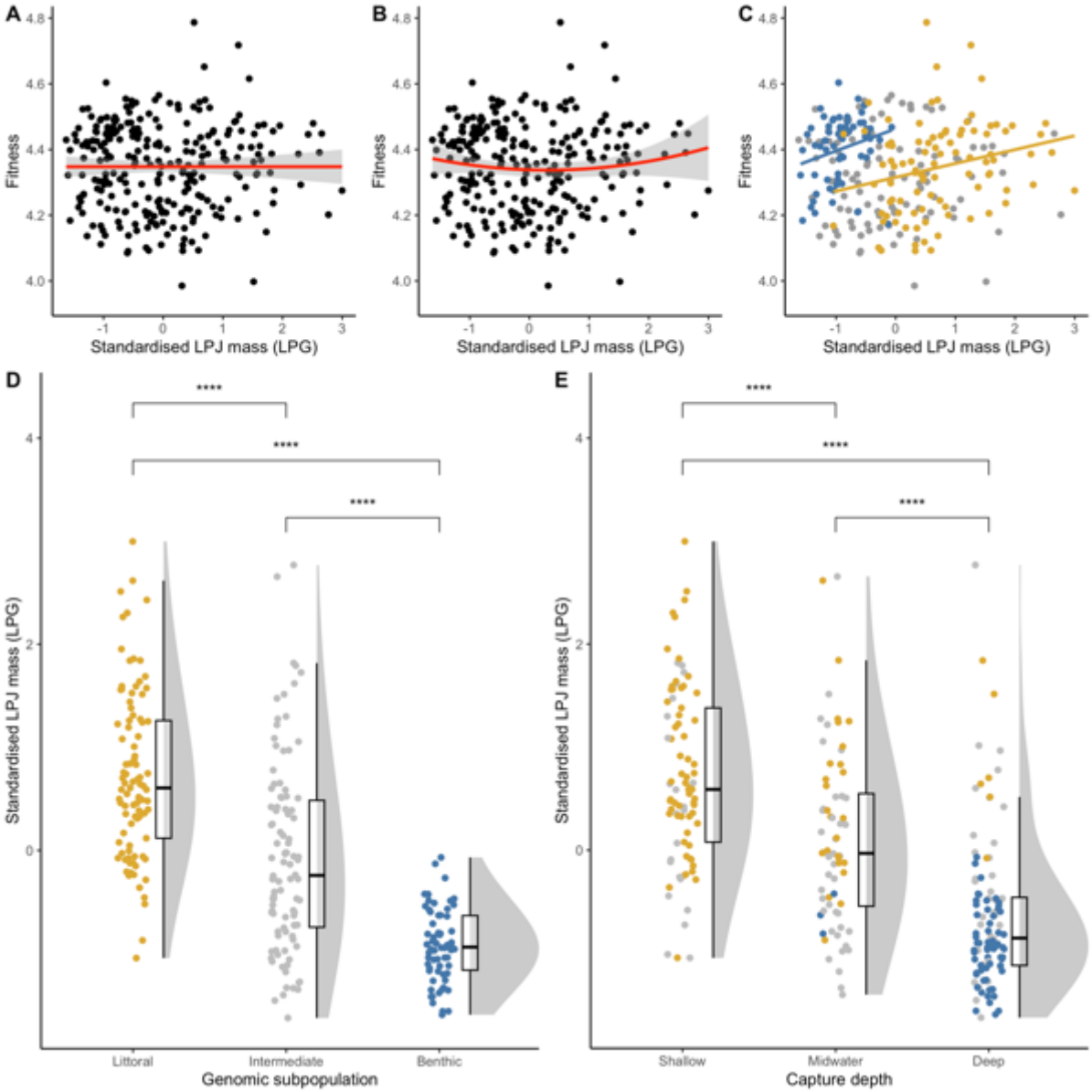


Figure S5.28. Standardised lower pharyngeal jaw (LPJ) mass in Lake Masoko. (A-C) Fitness surfaces for standardised LPJ mass, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised LPJ mass values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns p>0.05; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

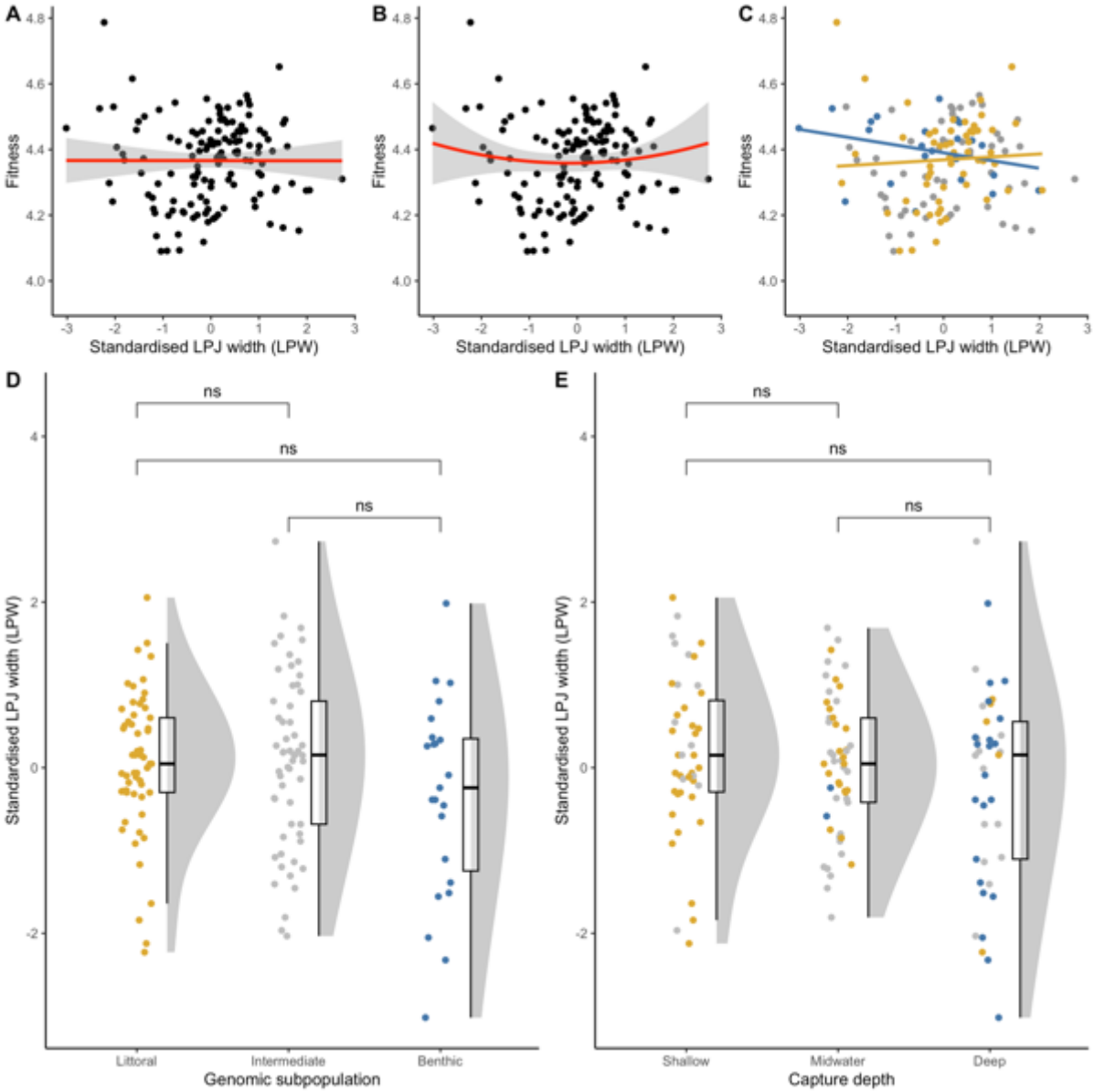


Figure S5.29. Standardised lower pharyngeal jaw (LPJ) width in Lake Masoko. (A-C) Fitness surfaces for standardised LPJ width, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised LPJ width values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns p>0.05; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

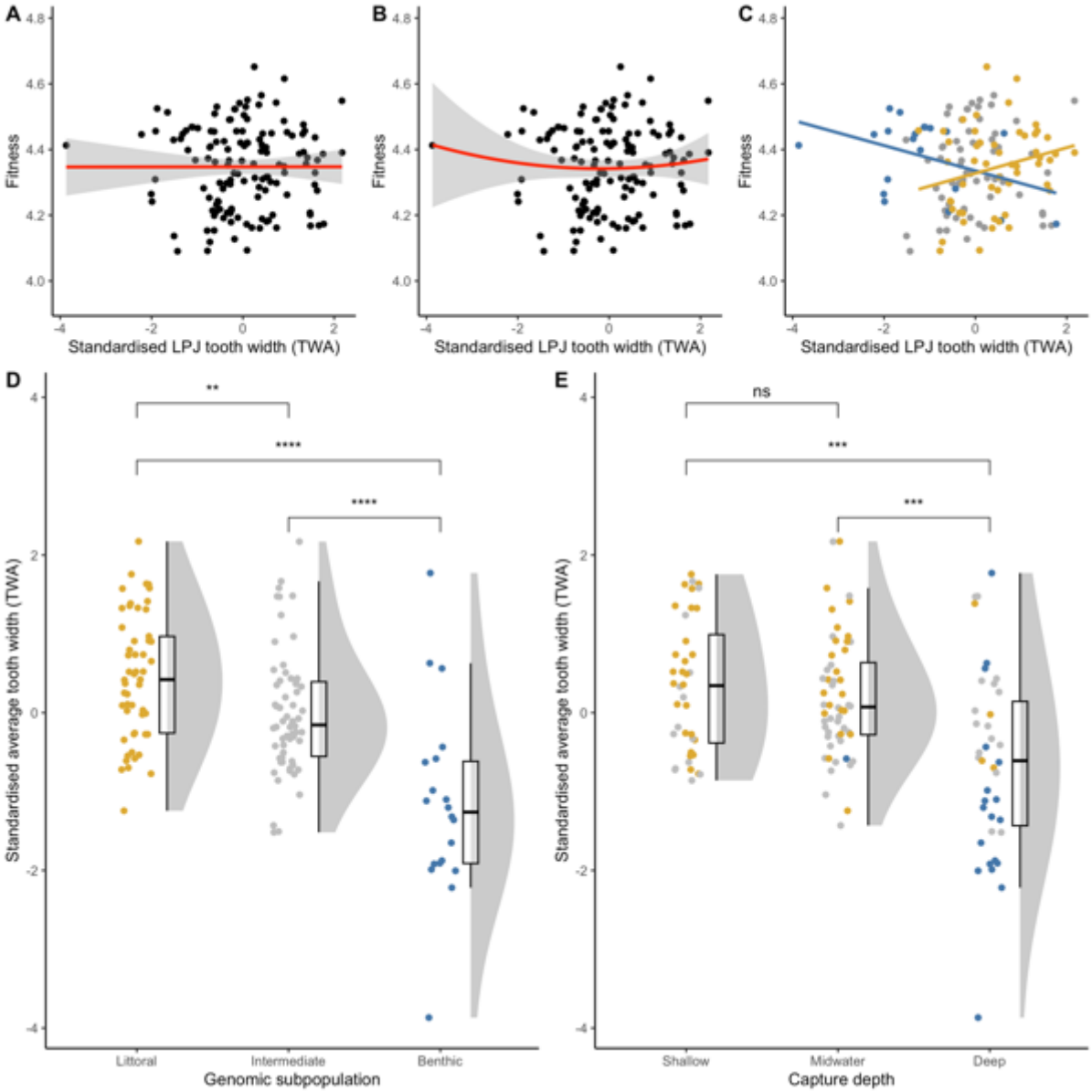


Figure S5.30. Standardised mean lower pharyngeal jaw (LPJ) tooth width in Lake Masoko. (A-C) Fitness surfaces for standardised mean LPJ tooth width, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised mean LPJ tooth width values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

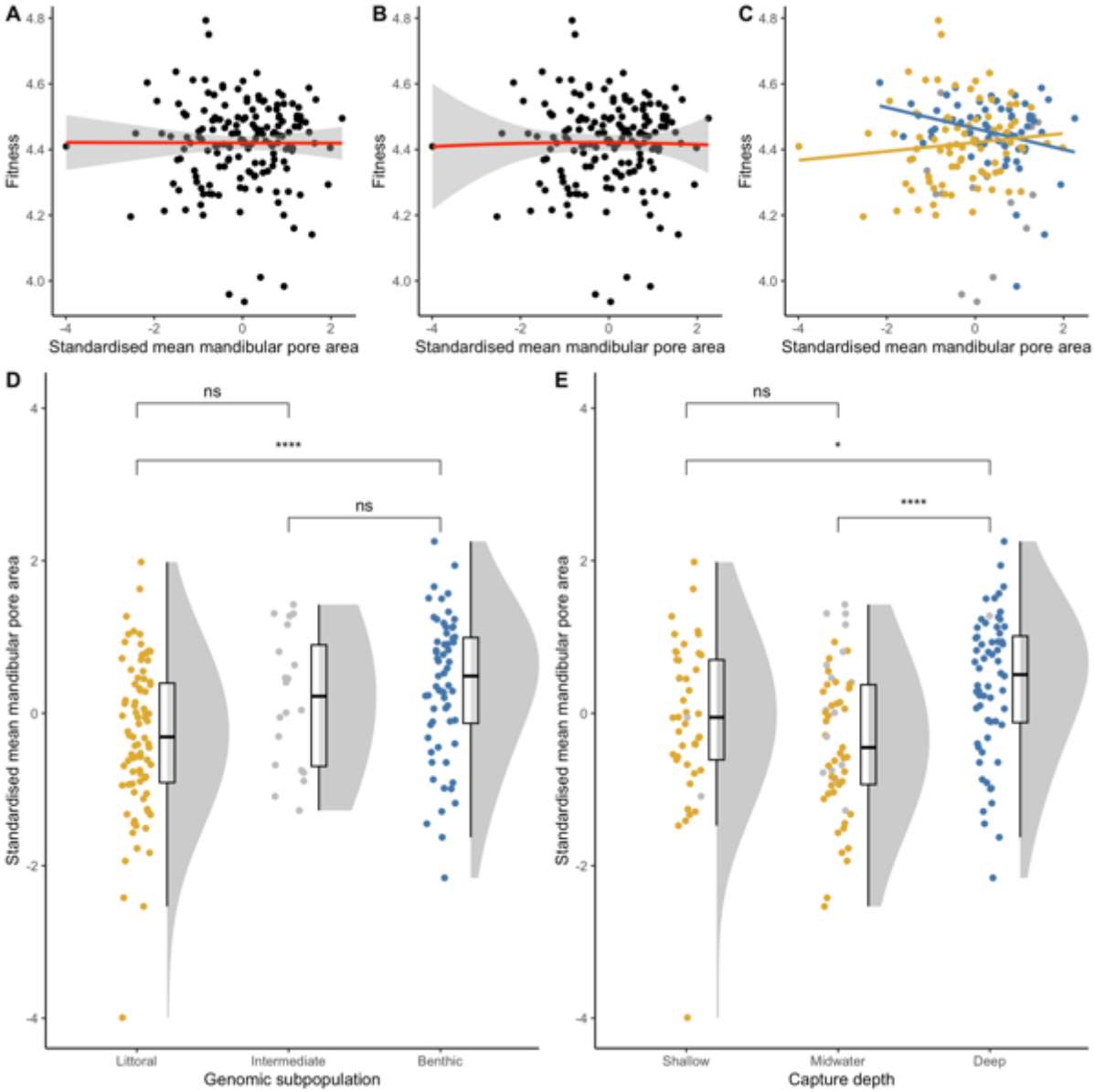


Figure S5.31. Standardised mean mandibular canal pore area in Lake Masoko. (A-C) Fitness surfaces for standardised mean mandibular canal pore area, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised mean mandibular canal pore area values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

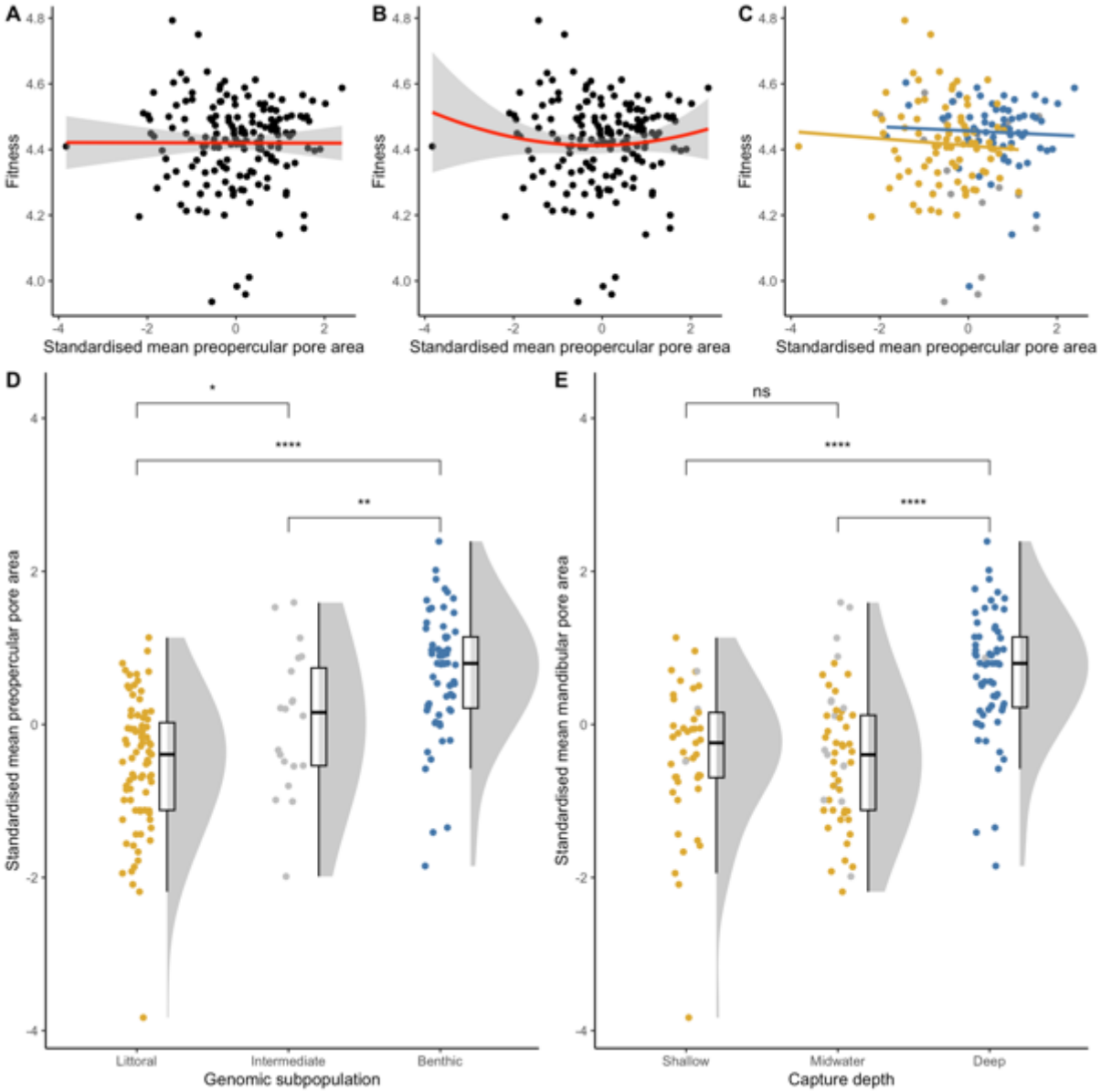


Figure S5.32. Standardised mean preopercular canal pore area in Lake Masoko. (A-C) Fitness surfaces for standardised mean preopercular canal pore area, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised mean preopercular canal pore area values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

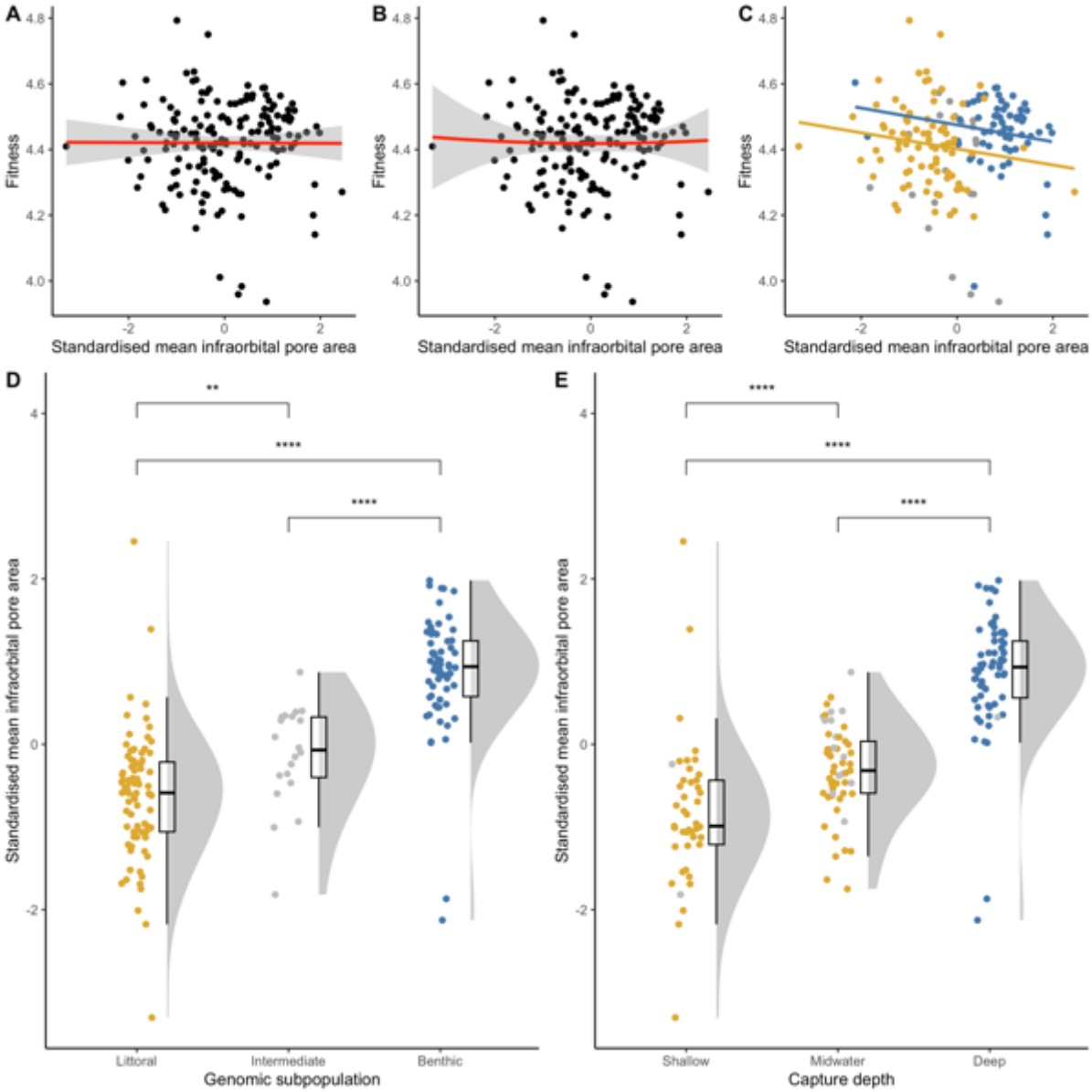


Figure S5.33. Standardised mean infraorbital canal pore area in Lake Masoko. (A-C) Fitness surfaces for standardised mean infraorbital canal pore area, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised mean infraorbital canal pore area values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

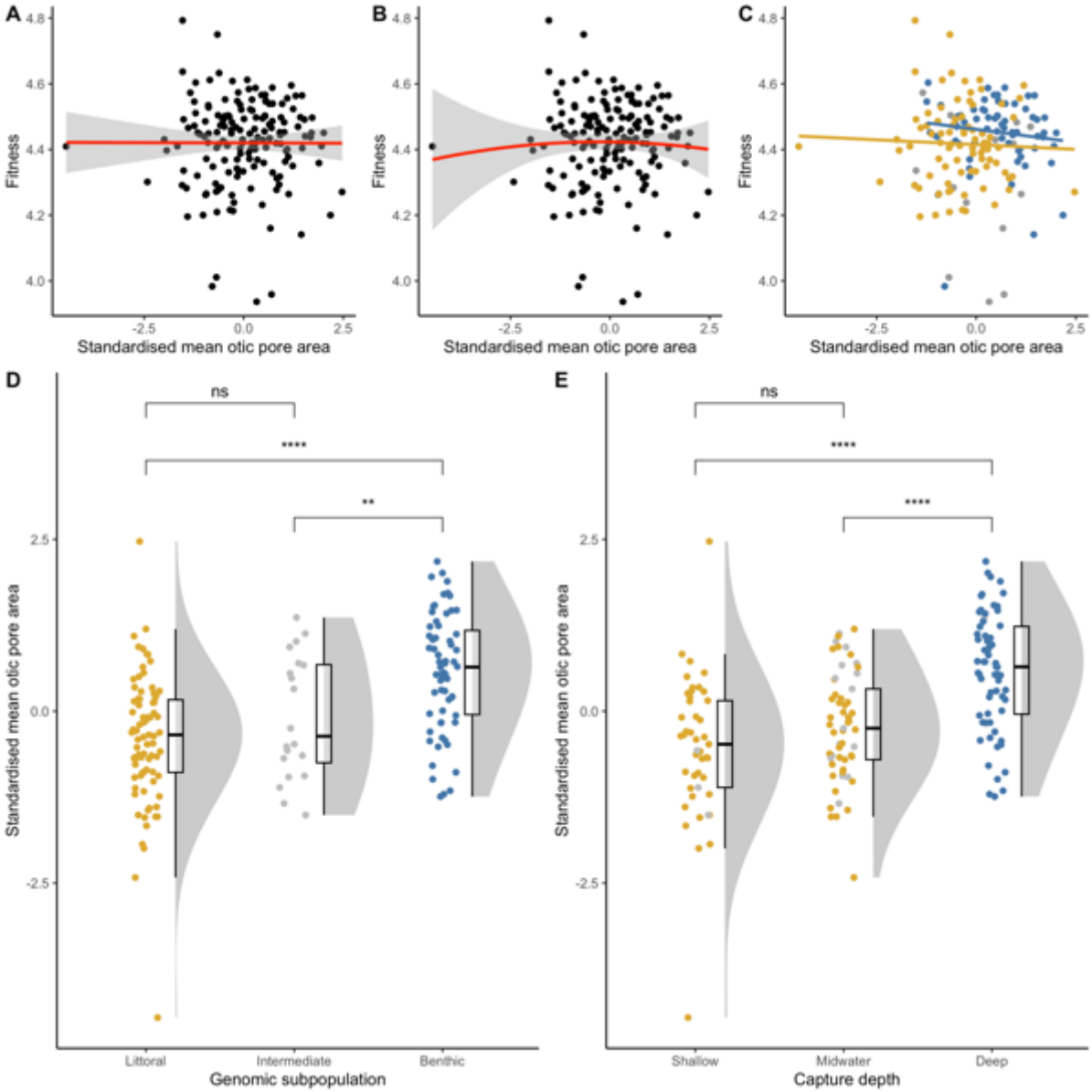


Figure S5.34. Standardised mean otic canal pore area in Lake Masoko. (A-C) Fitness surfaces for standardised mean otic canal pore area, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised mean otic canal pore area values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons

of a Kruskal-Wallis H test. ns p>0.05; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

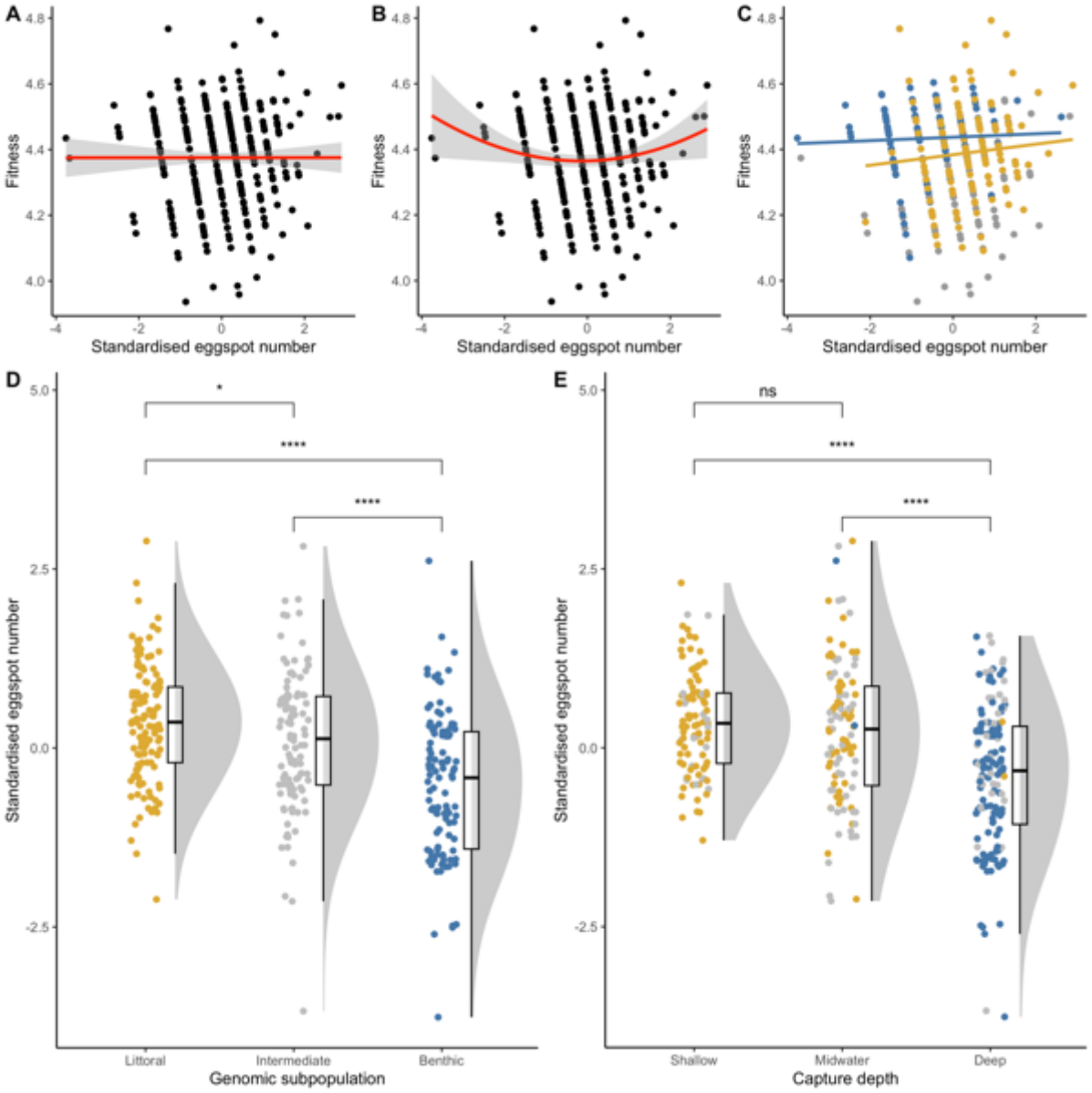


Figure S5.35. Standardised eggspot number in Lake Masoko. (A-C) Fitness surfaces for standardised eggspot number, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised eggspot number values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns p>0.05; *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

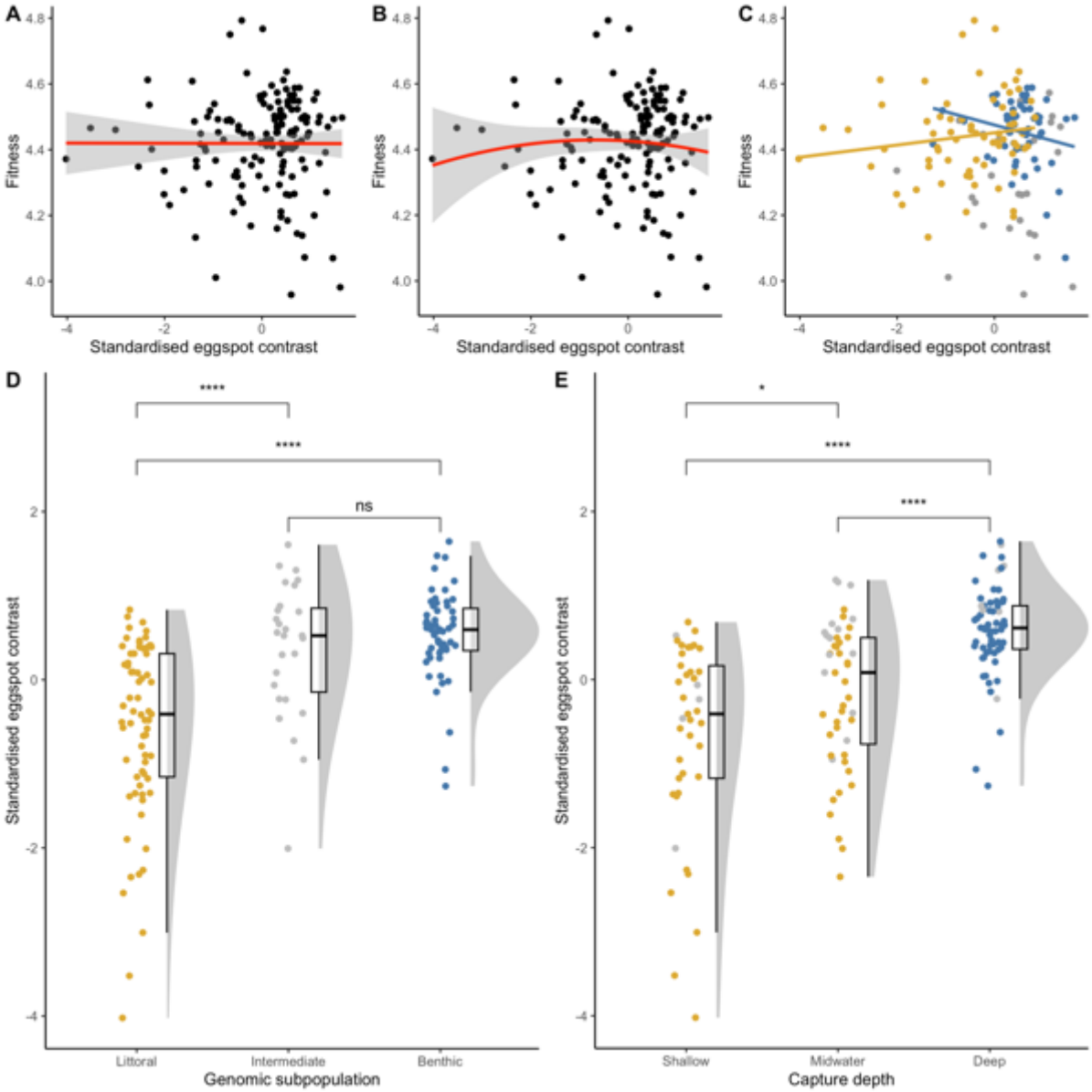


Figure S5.36. Standardised eggspot contrast in Lake Masoko. (A-C) Fitness surfaces for standardised eggspot contrast, with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised eggspot contrast values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

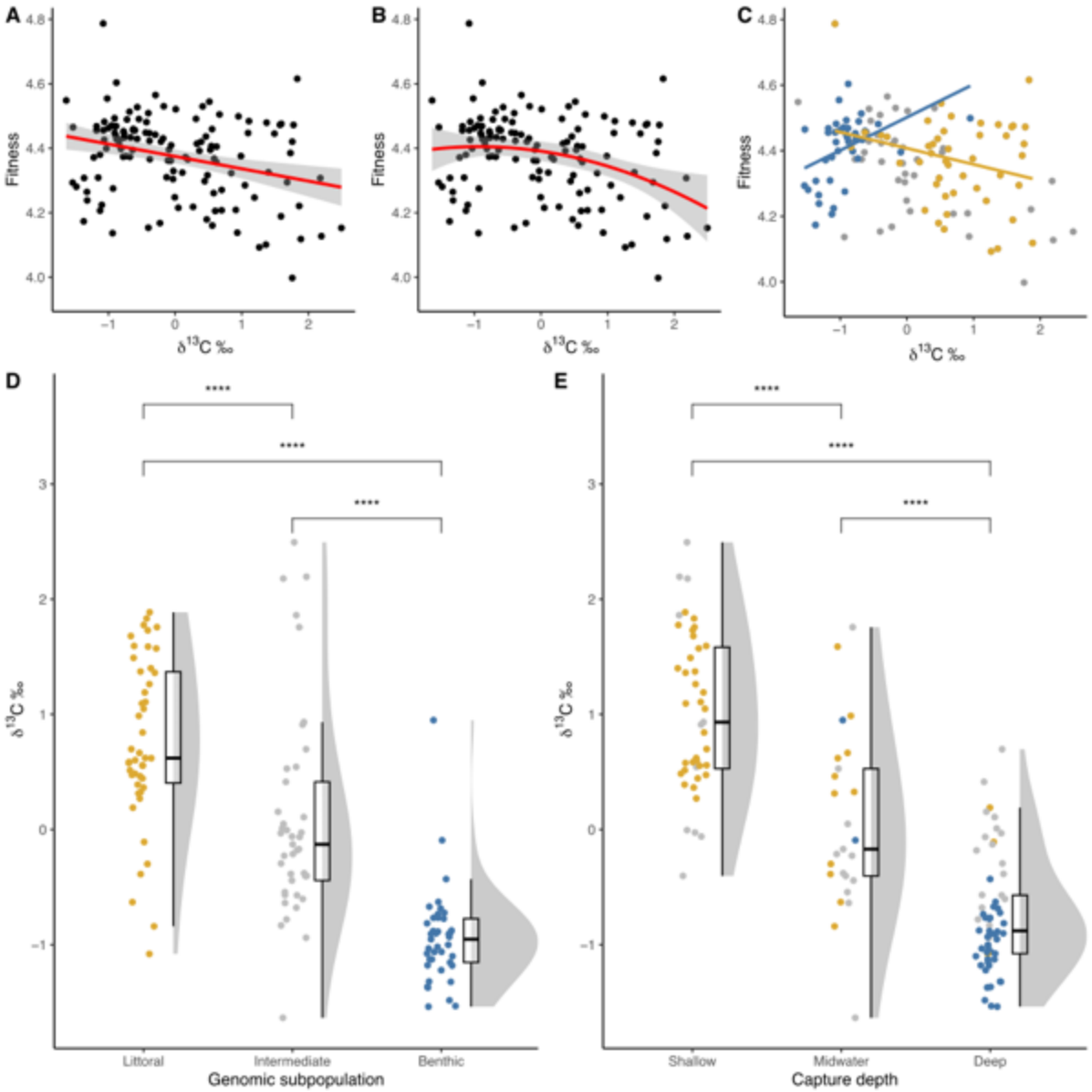


Figure S5.37. Standardised carbon isotope ratio (δ¹³C‰) in Lake Masoko. (A-C) Fitness surfaces for standardised carbon isotope ratio (δ¹³C‰), with (A) a whole population linear selection gradient, (B) whole population quadratic selection gradient, and (C) separate linear selection gradients for the littoral and benthic subpopulations. (D&E) Standardised carbon isotope ratio (δ¹³C‰) values for each (D) genomic subpopulation, and (E) capture depth, are shown. Comparisons between groups show pairwise comparisons of a Kruskal-Wallis H test. ns $p>0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

Sunset over Lake Tanganyika, Kigoma, August 2022
Photo: Duncan Edgley

Chapter 6

General Discussion

6.1 Overview

The cichlid adaptive radiations of the East African Great Lakes have long been of interest to evolutionary biologists (Seehausen, 2006; Seehausen, 2014), due to the rapid rate of evolutionary change (Malinsky et al. 2018; Meier et al. 2018), remarkable species richness (Seehausen, 2006), and most notably, the spectacular morphological diversity (Takahashi and Koblmüller, 2011). These radiations have become model systems for observing and testing evolutionary processes, particularly those related to the origin and function of morphological diversity (Seehausen, 2006). In most cases, the process of unpicking morphology's evolutionary origin in adaptive radiations initially involves describing the extent of variation present in traits across species (Albertson and Kocher, 2006; Takahashi and Koblmüller, 2011). But following this, an understanding of how these morphologies correlate with ecology and relate to function (e.g. Edgley and Genner, 2019; Konings et al. 2021), and how they arise genetically (Albertson and Kocher, 2006), are both needed to understand how species diversify. For example, in some cases morphology is more closely related to ecology than to phylogeny, implying it is functional for survival, and related to adaptation for species' ecological niche. This has led to insights into the function of morphological diversity in craniofacial shape (Albertson and Kocher, 2006), body shape (Husemann et al. 2014), body colouration (Albertson et al. 2014) and feeding morphology (Edgley and Genner, 2019), finding a spectacular diversity in form, in most cases linked with adaptive function.

The primary aim of the research presented in this thesis has been to add to this existing evidence on the adaptive function of morphological diversity in cichlids (Parsons et al. 2012; Maan and Sefc, 2013; Navon et al. 2017). I primarily focused on the mechanosensory lateral line system, which has been relatively poorly understudied from an evolutionary perspective (Mogdans, 2019). In the African cichlid fishes, little has been discovered about its diversity beyond a few individual species morphologies (Schwalbe et al. 2012; Schwalbe and Webb, 2014; Webb et al. 2014; Butler and Maruska, 2016) until recently (Edgley and Genner, 2019). This is despite it being a functionally important sensory modality for a range of behaviours in all fishes (Partridge and Pitcher, 1980; Satou et al. 1994; Montgomery et al. 1997; Coombs et al. 2001), including cichlids (Schwalbe et al. 2012; Butler and Maruska, 2015; Chapter 4).

Initially I built upon existing research describing morphological diversity among radiating cichlids in Lake Malawi (Edgley and Genner, 2019), and extended this to neuromast diversity (Chapter 2). Using a geometric morphometric approach, in Chapter 2 we aimed to determine whether lateral line system phenotypes are under divergent evolution in a system early in the speciation process, or whether this process occurs later during species divergence. Chapter 3 aimed to give insights into the genomic underpinning of lateral line system diversity through genotype-phenotype association analyses in an F_2 cross between species with divergent lateral line phenotypes. Similarly, in chapter 4, we aimed to determine the role of the lateral line system in mediating collective behaviour in cichlids, by subjecting F_2 individuals to behavioural trials, and correlating this with their lateral line system morphologies. Chapter 5 aimed to observe morphological diversification more broadly during population divergence, through measuring a broad spectrum of traits, and testing for strength of selection on phenotypes using a fitness proxy.

6.2 Lateral line diversification occurs during initial divergence along a depth cline

In Chapter 2, we characterise the diversity in lateral line system morphology in diversifying cichlid subpopulations of *Astatotilapia calliptera* in a small crater lake, Lake Masoko (Malinsky et al. 2015). We show that the emergence of pronounced differences in sensory morphology happens early during population divergence. We find that the morphology of the cephalic lateral line canals differed between ecomorphs within the Lake: specifically, cranial canal pores were significantly larger in the benthic than in the littoral ecomorph, indicating an increased reliance on flow stimuli in the deeper water (Figure 2.4). We found no evidence for subpopulation-linked differences in neuromast number in Lake Masoko (Figure 2.5). Although this lateral line system diversification has happened along a depth cline in short evolutionary time (~1,000 years) (Malinsky et al. 2015), the selection pressures causing this, and functional differences between these morphologies, are not known. Widened canal phenotypes (Figure 1.7B) are often hypothesised to be an adaptation for trophic niche (Schwable et al. 2012; Edgley and Genner, 2019). But it is often thought that such dietary niche specialisations happen later in cichlid adaptive radiations (Streelman and Danley, 2003; Ronco et al. 2021), following a broad initial adaptation for different depths (Seehausen, 2006; DeLorenzo et al. 2021).

To fully characterise the function of these neuroanatomical differences we observe, a better understanding of the ecology of these species is needed. For example, given that it may be the

case that these morphological modifications are adaptations to different trophic niches, evidence of dietary divergence between ecomorphs is required. Although stable isotope analysis has revealed changes in diet with depth in Lake Masoko *A. calliptera* (Figure 2.1E) (Carruthers et al. 2022), evidence for distributions of prey or food items in the Lake has only been quantified superficially (Malinsky et al. 2015; Turner et al. 2019). Ideally we would have a better understanding of relative abundances of prey across the Lake, and differences in mechaoreceptive prey-acquisition strategies (through behavioural trials or otherwise). The approach we took in Chapter 4, in correlating lateral line system phenotype with shoaling could be expanded for this purpose. Testing for behavioural differences in foraging efficiency/strategy when feeding on different prey types (Malinsky et al. 2015; Turner et al. 2019) and correlating this with morphology of aspects of the lateral line system, will help to determine the adaptive function of these structures. Beyond this, further investigation into the role of rate heterochrony and ontogeny in the generation of these differential phenotypes is required. Our findings that number of neuromasts (Webb 1989; Webb 2014b) and cephalic canal pore size (Bird and Webb, 2014) are associated with body size are consistent with the literature. In addition, it has been found that the differences between large-pored (*Aulonocara stuartgranti*) and small-pored species of cichlid (*Tramitichromis* sp.) can be in large part, attributed to differences in rate heterochrony (Bird and Webb, 2014). The developmental processes involved in generating these phenotypes may help to clarify their evolutionary origin and function (Montgomery et al. 1994; Nickles et al. 2020).

Further understanding of the function if lateral line diversification in Lake Masoko may be aided by determining the genomic basis for morphological disparity between ecomorphs. The initial evidence that the two ecomorphs in Masoko are in the early stages of speciation was derived from whole-genome sequence data, including observations that diversification has taken place in key functional genes related to speciation along a depth cline, such as the visual opsin gene rhodopsin, in the deeper and darker benthic environment (Malinsky et al. 2015). This existing whole-genome sequence data (outlined in Chapter 5) could be examined in light of our findings from Chapter 3, in which we propose candidate regions for genes involved in lateral line system diversity. If it is the case that Lake Masoko *A. calliptera* ecomorphs are diverging at loci we have found to be associated with lateral line canal pore size through an F_2 cross, this will serve to validate our findings in both analyses. Alternatively, we may find that these genes we have identified have differential expression profiles between ecomorphs (Vernaz et al., 2022; Carruthers et al. 2022). This may also be expanded to existing whole-genome sequence data or

gene expression data from Lake Malawi (Chapter 5; Malinsky et al. 2018). In addition, comparison with other systems, such as crater lake populations from related lineages which have also diversified along a depth cline (Moser et al. 2018; Barluenga et al. 2006), may allow us to determine whether the changes we observe here are unique to Lake Masoko, and whether the tempo and nature of morphological change occurs consistently and convergently across speciation events and during broader adaptive radiation (Meier et al. 2018; Ronco et al. 2021).

6.3 First insights into the genomic basis for lateral line system diversity

Our findings in chapter 3 give the first indications of genomic loci that may underpin variation in lateral line system diversity in cichlid fishes. We identified genes associated with GO terms enriched for craniofacial, neurological and developmental functions, that may be involved in the generation of phenotypic variation in our hybrid cross, but also across the parallel adaptive radiations of East Africa. Despite the initial evidence we show here, additional validation of genomic loci is required to determine whether candidates are truly associated with the lateral line system in cichlids. This may prove challenging, given the fact that such a complex system is likely to be highly polygenic (Wark et al. 2012; Greenwood et al. 2013). For example, our identification of the gene *bmp4*, a key gene in evolutionary biology, is notable, in that it has previously been linked to craniofacial and trophic morphology in cichlids and other lineages (Albertson et al. 2003; Abzhanov et al. 2004; Carruthers et al. 2022). However, *bmp4* is a broadly expressed and pleiotropic gene, so it may prove difficult to characterise its role in shaping lateral line system phenotypes through conventional methods (Abzhanov et al. 2003).

Furthermore, our candidate genes may be investigated in the wider adaptive radiation context. This may be possible by comparing our findings (Table 3.1) against existing genomic sequence data, or gene expression profiles from Lake Malawi (Malinsky et al. 2018), Lake Tanganyika (El Taher et al. 2020; Ronco et al. 2021) and Lake Victoria cichlids (Meier et al. 2018; Takuno et al. 2019), even in Neotropical cichlid lineages (Kautt et al. 2020). There may be signs of differential expression of these genes in species with highly divergent lateral line system phenotypes in these radiations. Comparing among lineages from different African radiations will determine whether the genomic underpinning we find in our cross is consistent using standing genetic variation (Matschiner et al. 2020; Masonick et al. 2022), or whether different genes and mutations are favoured under different circumstances. The use of CRISPR/Cas9 gene editing methods may help to determine the function of specific genes and genomic variants, an approach which has been utilised in cichlids previously with success, specifically with regard to patterning and

pigmentation-related loci (Kratchowil et al. 2018). Depending on how polygenic lateral line phenotypes are (Seleit et al. 2021), fine-scale mapping and validation of expression in tissues through in-situ hybridisation may be required. Knockout mutants at candidate loci may then be tested under laboratory conditions to determine behavioural consequences of loss of function, perhaps in shoaling behaviour (Chapter 4) or feeding efficiency. This validation of our candidate loci is particularly important, as it may be the case that genes are not actually associated with the lateral line system, but are expressed because of other factors, either directly or indirectly (Mehta et al. 2021; Carruthers et al. 2022). In addition to this, these genes may be linked with developmental processes, enabling insight into the ontogenetic nature of the origin of different lateral line system phenotypes, and the potential role of rate heterochrony in generation phenotypic disparity (Bird and Webb, 2014).

6.4 The lateral line system mediates cichlid collective behaviour

We found that some aspects of lateral line morphology, such as the number of neuromasts, are correlated with collective behaviour in cichlids (Chapter 4). The weak nature of these correlation perhaps reflects the broad spectrum of behaviours lateral line-mechanoreception informs, and that may be shaping its evolution (Mogdans, 2019). In cichlids for example, as well as informing shoaling/schooling behaviours that we illustrate here (Faucher et al. 2010; Chapter 4), the lateral line system is known to also mediate rheotactic detection of prey (Montgomery and McDonald, 1987; Montgomery, 1989; Schwalbe et al. 2012) and male-male territorial social interactions (Butler and Maruska, 2015) In addition, in one parent of our cross, *Aulonocara stuartgranti*, the widened cranial canal pore morphology is known to at least partially be an adaptation for the detection of buried prey (Schwalbe et al. 2012; Edgley and Genner, 2019). It may also be utilised in other behaviours too, but it seems that it is primarily a trophic adaptation in this species. Unpicking this complexity will require testing for many different behaviours in an F_2 cross, such as foraging efficiency and navigation in dark or turbid environments.

Importantly, our findings from a novel design of a lateral line sensor serve to validate our consideration of canal pore area as an important aspect of functional diversity. Increased pore sizes increased the sensitivity of an artificial sensor to detect vortices shed from a cylinder mimicking an upstream shoaling companion (Figure 4.5C). These findings help to inform interpretation of the results of our comparative morphology study in Lake Masoko (Chapter 2). For example, the larger pores found in the benthic ecomorph of *A. calliptera* in Lake Masoko, in increasing sensitivity to the detection of upstream companion fish, may serve to reduce the risk

of collision, and aid with navigation in the comparatively dark environment they inhabit (Malinsky et al. 2015; Turner et al. 2019).

Further investigation into our *Aulonocara stuartgranti* × *Otopharynx lithobates* F_2 cross here may also serve to validate our findings from chapter 3. The two crosses share one common parent, *A. stuartgranti*, and both have one different parent, *A. calliptera* (Chapter 3) and *O. lithobates* (Chapter 4). Given this, there is the future potential to genotype this *A. stuartgranti* × *O. lithobates* hybrid cross. Looking for overlap between genes identified through phenotype-genotype mapping in these two crosses may help to refine our mapping methodology, and determine which SNPs are truly associated with lateral line diversification. In this way, it can act as a confirmatory experiment to our findings in Chapter 3, and test the validity of our phenotype-genotype associations. Given our indications from whole-genome sequencing in Chapter 3, the genotyping of the Chapter 4 cross may be possible at lower sequencing depths, or perhaps employing targeted sequencing of specific genomic regions of interest (Davey et al. 2010; Albertson et al. 2014).

6.5 A novel approach to measuring and modelling disruptive selection

In Chapter 5, we used a proxy for relative fitness to measure natural selection in Lake Masoko cichlids across a range of genomic, morphological and ecological traits. We confirmed the existence of two broad subpopulations (or ecomorphs) associated with the benthic and littoral zones of the lake (Malinsky et al. 2015). When visualising the genomic and gross morphological adaptive landscapes, we found multiple fitness peaks or optima where the benthic and littoral subpopulations cluster, with intermediate or hybrid individuals inhabiting fitness valleys between them (Figure 5.2). In addition, when measuring natural selection on individual traits, we find evidence that conventional methods for inferring disruptive selection (Lande and Arnold, 1983; Kingsolver, 2001) are unable to capture the subtler patterns of divergence in a system such as Lake Masoko.

The Lake Masoko system provides a unique opportunity to measure natural selection, given our knowledge of population divergence between ecomorphs, and the existing phenotypic and genomic data (Malinsky et al. 2015). We propose that three kinds of disruptive selection are observed in Lake Masoko *A. calliptera*, all through differential trajectories of directional selection between subpopulations. These were not necessarily detectable using the conventional quadratic

selection gradients of Lande and Arnold (1983) but will nonetheless result in the accumulation of morphological differences over time as the populations diversify (Figure 5.5).

In the light of our findings, mathematical modelling of evolutionary change in Lake Masoko may be utilised (Beaumont, 2010) to predict future evolutionary change in Lake Masoko. Through an expansion of the methods we outline here, it may be possible to develop a framework for measuring divergence in traits between populations, which incorporates our methods with those outlined in Lande and Arnold's seminal paper (1983). This may involve developing a parameter, analogous to the selection gradient or selection differential, which describes and quantifies type I, II and III disruptive selection as we define them here (related to a GLM/ANCOVA interaction term or otherwise). This parameter will need to take into account correlation between traits, and its accuracy will be highly dependent on accurate characterisation of what defines subpopulations, either genomic or phenotypic. It may need to incorporate some measure of the degree of reproductive isolation/gene flow between populations into its calculation (Malinsky et al. 2021). Our definition of disruptive selection may also be expanded to consider more complex systems, such as those occurring in the African cichlid adaptive radiations, in which population divergence in not necessarily occurring in a binary fashion between deep and shallow; it may involve a complex web of semi-isolated populations, all of which exhibit differing trajectories of selection on morphology, and with differing degrees of gene flow between them. Modelling selection in such a system will prove challenging in its complexity, but may give important insights into how populations diversify and isolate into different species during adaptive radiation.

One primary concern when measuring natural selection, is the validation of the choice of fitness proxy (Hoekstra et al. 2001). Condition based fitness proxies such as we employ here, have been used in the literature (Franklin and Morrissey, 2017; reviewed in Viblanc et al. 2022), but have many associated problems depending on experimental design, and their application during statistical analysis, even in cases of high sample size, such that we have here (Viblanc et al. 2022). Standard length is highly correlated with reproductive success and survival in haplochromine cichlids, however there may be many other factors causing variation in size beyond the scope of straightforward evolutionary fitness.

Given these issues, it may be possible to measure fitness in fishes' natural environments more accurately, and in a more quantitative way. For example, it may be possible simply to look at

reproductive success directly, using cage-based setups *in situ*. Alternatively, it may be possible to measure survival rate as a fitness measure, as originally proposed (Lande and Arnold, 1983). This will prove challenging, given the primary cause of mortality in wild Lake Masoko *A. calliptera*, is predation (Turner et al. 2019). Existing studies that consider survival, often involve translocation of fish to a new environment, and genotyping/phenotyping survivors, in a mark-recapture setup (Calsbeek and Smith, 2008; Hendry et al. 2009; Schluter et al. 2021). Composite condition-based fitness measures are also possible, but require more effort to measure, are less immediately comprehensible, and are less verifiable (Martin and Wainwright, 2013).

Here, it is important to clarify whether age-based proxies (standard length, scale circulus count) such as ours are actually related to reproductive success. In haplochromine cichlids, it is known that larger males have more mating success (Markert and Arnegard 2007), but this is not the case in other lineages, and is not known specifically for male *A. calliptera,* neither from their natural environment nor under aquarium conditions. This may be determined in aquaria through mate choice experiments, and testing for correlations between the number of successful matings and standard length, perhaps through parentage analysis (Vandeputte and Haffray, 2014). Using an aquaria-based setup that mimics the conditions of the natural environments of the two ecomorphs, it may be possible to track relative fitness, and hence measure natural selection over time in controlled conditions. For example, the O_2 concentration, temperature, light level and diet can all be controlled in aquaria (Genner et al. preliminary results), and mimic that of the wild. Taking this approach, a more accurate measure of how selection operates on traits over time may be able to be established, given what we present here is merely a snapshot in time.

Our methods revealed no evidence for disruptive selection in Lake Masoko on pore size in any of the four cephalic lateral line canals we investigated (Table 5.1). This runs contrary to our findings in Chapter 2, in which the two have significantly different cephalic lateral line canal phenotypes (Figure 2.4), which we propose is a result of divergent selection, given different mechanoreception requirements in different depths in the lake. This may indicate that divergence has already occurred and stabilised at its current extent in the short time in which the populations have achieved partial reproductive isolation (~1,000 years). It may alternatively be the case that the range of possible phenotypes in the benthic ecomorph is constrained, perhaps by sequential selective sweeps (Malinsky et al. 2015). This is unlikely, however, given that the range of phenotypes among benthic individuals extends beyond the maximum range that we observe among littoral individuals (Figure 2.4), suggesting that it is not merely constraint in

phenotype or plasticity causing this disparity, but a novel evolutionary change. Thus, it will probably have a distinct genomic signature, which may be detectable, particularly in light the initial insights into this we show here (Chapter 3).

6.6 General conclusions

Critically, we show here that understanding how fishes adapt to their environments during evolution requires a consideration of mechanoreception via the lateral line. This is particularly pertinent in habitats where the use of other sensory modalities such as vision, is reduced, such as darker or more turbid waters (Yoshizawa et al. 2010; Yoshizawa et al. 2014), or for niches involving feeding on cryptic prey (Schwable et al. 2012), both of which are extremely common across the adaptive radiations of East Africa. Despite this, further research is required in order to determine the system's specific role in facilitating ecological adaptation. This will require a more thorough quantification of morphological disparity, including in the parallel adaptive radiations of Lake Victoria and Lake Tanganyika, which have not been investigated, despite the existence of high quality genomic and microCT data (Meier et al. 2018; Ronco et al. 2021). Additional behavioural trials linking phenotype to function can serve to unpick the selection pressures driving morphological diversification. These may be expanded to alternative behaviours and additional lateral line system phenotypes. Furthermore, finding the genomic regions involved in generating variation in lateral line system structures will help to determine the fundamental evolutionary processes underpinning observed variation, and whether this is consistent across parallel radiations.

Despite our findings, it remains extremely difficult to unpick the specific functions of aspects of morphological diversity, particularly a system as complex as the lateral line system (Mogdans, 2019). Here, we have highlighted broad patterns of morphological change, showing that the system exhibits niche-specific variation early in the diversification process (Chapter 2), and that variation can be correlated with individual diet (Edgley and Genner, 2019), and collective behaviour (Chapter 4). Further research is now needed to unpick the evolutionary role of the system in niche specialisation during adaptive radiations, and this will require a range of approaches, including the integration of behavioural, comparative morphology, and genomic methods. We suggest that future research considers the lateral line system as an intrinsic part of cichlid functional morphology, and as yet another example of the spectacular morphological diversity that arises during cichlid adaptive radiation. We hope that the results described in this

thesis provide a strong basis on which to build further knowledge and understanding of the fish lateral line system.

Reference list

Abdel-Latif, H., Hassan, E.S., and von Campenhausen, C. (1990) Sensory performance of blind Mexican cave fish after destruction of the canal neuromasts. *Die Naturwissenschaften,* 77(5), 237–239.

Abzhanov, A., Protas, M., Grant, B.R., Grant, P.R., and Tabin, C.J. (2004) Bmp4 and morphological variation of beaks in Darwin's finches'. *Science,* 305(5689): 1462–1465.

Adams, D., Collyer, M., Kaliontzopoulou, A., and Baken, E. (2022) geomorph: Software for geometric morphometric analyses. R package version 3.0.5. https://cran.r-project.org/package=geomorph.

Ahi, E.P., Singh, P., Duenser, A., Gessl, W., and Sturmbauer, C. (2019) Divergence in larval jaw gene expression reflects differential trophic adaptation in haplochromine cichlids prior to foraging. *BMC Evolutionary Biology,* 19(1): 150.

Ahnelt, H., Ramler, D., Madsen, M.Ø., Jensen, L.F., and Windhager, S. (2021) Diversity and sexual dimorphism in the head lateral line system in North Sea populations of threespine sticklebacks, Gasterosteus aculeatus (Teleostei: Gasterosteidae). *Zoomorphology,* 140(1): 103–117.

Akanyeti, O., Thornycroft, P.J.M., Lauder, G.V., Yanagitsuru, Y.R., Peterson, A.N., and Liao, J.C. (2016) Fish optimize sensing and respiration during undulatory swimming. *Nature Communications*, 7(1): 1-8.

Albertson, R. C., and Kocher, T.D. (2006) Genetic and developmental basis of cichlid trophic diversity. *Heredity,* 97(3): 211–21.

Albertson, R. C., J. T. Streelman, and T. D. Kocher. (2003) Genetic basis of adaptive shape differences in the cichlid head. *Journal of Heredity,* 94(4): 291–301.

Albertson, R.C., and Kocher, T.D. (2005) Genetic architecture sets limits on transgressive segregation in hybrid cichlid fishes. *Evolution,* 59(3): 686–690.

Albertson, R.C., Powder, K.E., Hu, Y., Coyle, K.P., Roberts, R.B., and Parsons, K.J. (2014) Genetic basis of continuous variation in the levels and modular inheritance of pigmentation in cichlid fishes. *Molecular Ecology,* 23(21): 5135–5150.

Albertson, R.C., Kawasaki, K.C., Tetrault, E.R., and Powder, K.E. (2018) Genetic analyses in Lake Malawi cichlids identify new roles for Fgf signaling in scale shape variation. *Communications Biology,* 1(1): 1–11.

Alexa, A., and Rahnenfuhrer, J. (2022) TopGO: enrichment analysis for gene ontology. Bioconductor version: Release 3.16. https://bioconductor.org/packages/release/bioc/html/topGO.html

Archambeault, S.L., Bärtschi, L.R., Merminod, A.D., and Peichel, C.L. (2020) Adaptation via pleiotropy and linkage: association mapping reveals a complex genetic architecture within the Stickleback Eda locus'. *Evolution Letters,* 4(4): 282–301.

Attwell, J.R., Ioannou, C.C., Reid, C.R., Herbert-Read, J.E. (2021) Fish avoid visually noisy environments where prey targeting is reduced. *American Naturalist,* 198(3): 421–432.

Bak-Coleman, J., Court, A.D., Paley, A., and Coombs, S. (2013) The spatiotemporal dynamics of rheotactic behavior depends on flow speed and available sensory information. *Journal of Experimental Biology,* 216(21): 4011–4024.

Baker, C.F., and Montgomery, J.C. (1999) Lateral line mediated rheotaxis in the Antarctic fish *Pagothenia borchgrevinki. Polar Biology,* 21(5): 305-309.

Baker, C.V.H., Modrell M.S., and Gillis, J.A. (2013) The evolution and development of vertebrate lateral line electroreceptors. *The Journal of Experimental Biology,* 216(13): 2515–2522.

Barker, P., Williamson, D., Gasse, F., Gibert, E. (2003) Climatic and volcanic forcing revealed in a 50,000-year diatom record from Lake Massoko, Tanzania. *Quaternary Research,* 60(3): 368–376.

Barluenga, M., Stölting, K.N., Salzburger W., Muschick, M., and Meyer, A. (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature,* 439(7077): 719–723.

Barneche, D.R., Robertson, D.R., White, C.R., and Marshall, D.J. (2018) Fish reproductive-energy output increases disproportionately with body size. *Science,* 360(6389): 642–645.

Baumgarten, L., Machado-Schiaffino, G., Henning, F., and Meyer, A. (2015) What big lips are good for: on the adaptive function of repeatedly evolved hypertrophied lips of cichlid fishes. *Biological Journal of the Linnean Society,* 115(2): 448–455

Bassett, D.K., Carton, A.G., and Montgomery, J.C. (2006) Flowing water decreases hydrodynamic signal detection in a fish with an epidermal lateral-line system. *Marine and Freshwater Research,* 57(6): 611–617.

Beaumont, M.A. (2010) Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41: 379–406.

Beausoleil, M., Frishkoff, L.O., M'Gonigle, L.K., Raeymaekers, J.A.M., Knutie, S.A., De León, L.F., Huber, S.K. et al. (2019) Temporally varying disruptive selection in the medium ground finch (Geospiza fortis). *Proceedings of the Royal Society B: Biological Sciences,* 286(1916): 20192290.

Becker, E.A., Bird, N.C., and Webb, J.F. (2016) Post-embryonic development of canal and superficial neuromasts and the generation of two cranial lateral line phenotypes. *Journal of Morphology,* 277(10): 1273–1291.

Beckmann, M., Erős, T., Schmitz, A., and Bleckmann, H. (2010) Number and distribution of superficial neuromasts in twelve common European Cypriniform fishes and their relationship to habitat occurrence. *International Review of Hydrobiology,* 95(3): 273–284.

Bell, M.A., and Travis, M.P. (2005) Hybridization, transgressive segregation, genetic covariation, and adaptive radiation. *Trends in Ecology & Evolution,* 20(7): 358–361.

Benkman, C.W. (2003) Divergent selection drives the adaptive radiation of crossbills. *Evolution,* 57(5): 1176–1181.

Bereiter-Hahn, J. (1976) Dimethylaminostyrylmethylpyridiniumiodine (daspmi) as a fluorescent probe for mitochondria in situ. *Biochimica et Biophysica Acta,* 423(1): 1–14.

Bird, N.C., and Webb, J.F. (2014) Heterochrony, modularity, and the functional evolution of the mechanosensory lateral line canal system of fishes. *EvoDevo,* 5(1): 21.

Blaxter, J.H.S., and Batty, R.S. (1985) Herring behaviour in the dark: responses to stationary and continuously vibrating obstacles. *Journal of the Marine Biological Association of the UK,* 65: 1031–1049.

Bleckmann, H. (1988) Prey identification and prey localization in surface-feeding fish and fishing spiders. In Atema, J., Fay, R.R., Popper, A.N., and Tavolga, W.N. (Eds.), *Sensory biology of aquatic animals*. Springer, New York, NY. 619-641

Bleckmann, H., and Zelick, R. (2009) Lateral line system of fish. *Integrative Zoology,* 4(1): 13–25.

Bleckmann, H., Breithaupt, T., Blickhan, R., and Tautz, J. (1991) The time course and frequency content of hydrodynamic events caused by moving fish, frogs, and crustaceans. *Journal of Comparative Physiology A,* 168(6): 749–757.

Bleckmann, H., Tittel, G., Blübaum-Gronau, E. (1989) The lateral line system of surface-feeding fish: Anatomy, physiology, and behavior. In Coombs, S., Görner, P., Münz, H. (Eds.), *The Mechanosensory Lateral Line*. Springer, New York, NY, 501–526.

Bolker, B., Giné-Vázquez, I., and R Development Core Team (2017). bbmle: Tools for General Maximum Likelihood Estimation. R package version 1.0.20. https://CRAN.R-project.org/package=bbmle.

Bolnick, D.I., and Fitzpatrick, B.M. (2007) Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics,* 38(1): 459–487.

Braun, C.B., and Northcutt, R.G. (1997) The lateral line system of hagfishes (Craniata: Myxinoidea). *Acta Zoologica,* 78(3): 247–268.

Brionne, A., Juanchich, A., and Hennequet-Antier, C. (2019) ViSEAGO: a bioconductor package for clustering biological functions using gene ontology and semantic similarity. *BioData Mining,* 12(1): 16. https://bioconductor.org/packages/release/bioc/html/ViSEAGO.html

Brodie, E.D., Moore, A.J., and Janzen, F.J. (1995) Visualizing and quantifying natural selection. *Trends in Ecology & Evolution,* 10(8): 313–318.

Brown, A.D., Mussen, T.D., Sisneros, J.A., and Coffin, A.B. (2011) Reevaluating the use of aminoglycoside antibiotics in behavioral studies of the lateral line. *Hearing Research,* 272(1–2): 1–4

Brown, E.E.A., and Simmons, A.M. (2016) Variability of rheotaxis behaviors in larval bullfrogs highlights species diversity in lateral line function. *PLOS One,,* 11(11): e0166989.

Brumm, H. (2013) *Animal Communication and Noise, Animal Signals and Communication*. Springer, Berlin, Heidelberg.

Buck, L.M.J., Winter, M.J., Redfern, W.S., and Whitfield, T.T. (2012) Ototoxin-induced cellular damage in neuromasts disrupts lateral line function in larval zebrafish. *Hearing Research,* 284(1–2): 67–81.

Burke, M.K. (2012) How does adaptation sweep through the genome? insights from long-term selection experiments. *Proceedings of the Royal Society B: Biological Sciences,* 279(1749): 5029–5038.

Burress, E.D. (2016) Ecological diversification associated with the pharyngeal jaw diversity of neotropical cichlid fishes. *Journal of Animal Ecology,* 85(1): 302–13.

Butler, J.M., and Maruska, K.P. (2015) The mechanosensory lateral line is used to assess opponents and mediate aggressive behaviors during territorial interactions in an African cichlid fish. *The Journal of Experimental Biology,* 218(20): 3284–3294.

Butler, J.M., and Maruska, K.P. (2016) Mechanosensory signaling as a potential mode of communication during social interactions in fishes. *Journal of Experimental Biology,* 219(18): 2781–2789.

Butler, J.M., Field, K.E. and Maruska, K.P. (2016) Cobalt chloride treatment used to ablate the lateral line system also impairs the olfactory system in three freshwater fishes. *PLOS One,* 11(7): e0159521.

Butler, M.A., Sawyer, S.A., and Losos, J.B. (2007) Sexual dimorphism and adaptive radiation in Anolis lizards. *Nature,* 447(7141): 202–205.

Calsbeek, R., and Smith, T.B. (2008) Experimentally replicated disruptive selection on performance traits in a Caribbean lizard. *Evolution,* 62(2): 478–484.

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

Carruthers, M., Edgley, D.E., Saxon, A.D., Gabagambi, N.P. Shechonge, A., Miska, E.A., Durbin, R., Bridle, J.R., Turner, G.F., and Genner, M.J. (2022) Ecological speciation promoted by divergent regulation of functional genes within African cichlid fishes. *Molecular Biology and Evolution*, 39(11): msac251.

Carton, A.G., and Montgomery, J.C. (2004) A comparison of lateral line morphology of blue cod and torrentfish: two sandperches of the family Pinguipedidae. *Environmental Biology of Fishes,* 70(2): 123–131.

Chamberlain, A.C., and Ioannou, C.C. (2019) Turbidity increases risk perception but constrains collective behaviour during foraging by fish shoals. *Animal Behaviour,* 156: 129–138.

Chang, C.C., Chow, C.C., Laurent, C.T., Vattikuti, S., Purcell, S.M., and Lee J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience,* 4(1): 7. https://www.cog-genomics.org/plink/2.0/

Chiu, L.L., Cunningham, L.L., Raible, D.W., Rubel, E.W., and Ou, H.C. (2008) Using the zebrafish lateral line to screen for ototoxicity. *Journal of the Association for Research in Otolaryngology,* 9(2): 178–190.

Chivers, D.P., and Smith, R.J.F. (1998) Chemical alarm signalling in aquatic predator-prey systems: a review and prospectus. *Écoscience,* 5(3): 338–352.

Clark, B., Elkin, J., Marconi, A., Turner, G.F., Smith, A.M., Joyce, D., Miska, E.A., Juntti, S.A., and Santos, M.E. (2022) Oca2 targeting using CRISPR/Cas9 in the Malawi cichlid *Astatotilapia calliptera*. *Royal Society Open Science,* 9(4): 220077.

Conith, A.J., and Albertson, R.C. (2021) The cichlid oral and pharyngeal jaws are evolutionarily and genetically coupled. *Nature Communications,* 12(1): 5477.

Coombs, S. and Montgomery, J.C. (1994) Function and evolution of superficial neuromasts in an Antarctic notothenioid fish. *Brain, Behavior and Evolution,* 44(6): 287–298.

Coombs, S., and Montgomery, J.C. (1992) Fibers innervating different parts of the lateral line system of an Antarctic notothenioid, Trematomus bernacchii, have similar frequency responses, despite large variation in the peripheral morphology. *Brain, Behavior and Evolution,* 40(5): 217–233.

Coombs, S., Bak-Coleman, J., and Montgomery, J.C. (2020) Rheotaxis revisited: a multi-behavioral and multisensory perspective on how fish orient to flow. *Journal of Experimental Biology,* 223(23): jeb223008

Coombs, S., Braun, C.B., and Donovan, B. (2001) The orienting response of Lake Michigan mottled sculpin is mediated by canal neuromasts. *Journal of Experimental Biology,* 204(2): 337–348.

Coombs, S., Görner, P. and Münz, H. (1989) A brief overview of the mechanosensory lateral line system and the contributions to this volume. In Coombs, S., Görner, P., and Münz, H. (Eds.), *The Mechanosensory Lateral Line*, Springer, New York, Springer. 3–5.

Coombs, S., Janssen, J., and Webb, J.F. (1988) Diversity of lateral systems: evolutionary and functional considerations. In Atema, J., Fay, R.R., Popper, A.N., Tavolga, W.N. (Eds.) *Sensory Biology of Aquatic Animals*, Springer, New York, NY, 553–593.

Cooper, W.J., Parsons, K., McIntyre, A., Kern, B., McGee-Moore, A., and Albertson, R.C. (2010) Bentho-pelagic divergence of cichlid feeding architecture was prodigious and consistent during multiple adaptive radiations within African rift-lakes. *PLOS One,* 5(3): e9551.

Coyne, J.A., and Orr, H.A. (2004) *Speciation*. Oxford University Press, Oxford, UK.

Czech-Damal., N.U., Liebschner, A., Miersch, L., Klauer, G., Hanke, F.D., Marshall, C., Dehnhardt, G., and Hanke, W., Electroreception in the Guiana dolphin (*Sotalia guianensis*). *Proceedings of the Royal Society B: Biological Sciences,* 22(279): 663–668.

Danley, P.D., and Kocher, T.D. (2001) Speciation in rapidly diverging systems: lessons from Lake Malawi. *Molecular Ecology,* 10(5): 1075–1086.

Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life*. London, UK, John Murray.

Darwin, C., and Wallace, A.R. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Zoological Journal of the Linnean Society,* 3(9): 45–62.

Davey, J.W., and Blaxter, M.L. (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, 9(5-6): 416–423.

Davies, R.W., Flint, J., Myers, S., and Mott, R. (2016) Rapid genotype imputation from sequence without reference panels. *Nature Genetics,* 48(8): 965–969.

de Visser, A.G.M., and Krug, J. (2014) Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics,* 15(7): 480–490.

De-Kayne, R., Selz, O.M., Marques, D.A., Frei, D., Seehausen, O., and Feulner, P.G.D. (2022) Genomic architecture of adaptive radiation and hybridization in Alpine whitefish. *Nature Communications*, 13(1): 4479.

Delalande M. (2008) Hydrologie et geochimie isotopique du Lac Masoko et de lacs volcaniques de la province active du Rungwe (sud-ouest Tanzanie). PhD Thesis, Université Paris Sud – Paris.

DeLorenzo, L., Mathews, D., Brandon, A.A., Joglekar, M., Baez, A.C., Moore, E.C., Ciccotto, P.J., Roberts, N.B., Roberts, R.B., and Powder, K.E. (2021) Genetic basis of body shape variation along the benthic-pelagic axis in cichlid fishes. bioRxiv. https://doi.org/10.1101/2021.10.02.462884.

Dijkgraaf, S. (1963) The functioning and significance of the lateral-line organs. *Biological Reviews*. 38: 51-105.

Dijkgraaf, S., and Kalmijn, A.J. (1963) Untersuchungen über die funktion der lorenzinischen ampullen an haifischen. *Zeitschrift für vergleichende Physiologie,* 47(4): 438–456.

Doebeli, M., and Dieckmann, U. (2003) Speciation along environmental gradients. *Nature,* 421(6920): 259–264.

Doyle, R.W., Talbot, A.J., and Nicholas, R.R. (1987) Statistical interrelation of length, growth, and scale circulus spacing: appraisal of a growth rate estimator for fish. *Canadian Journal of Fisheries and Aquatic Sciences,* 44(9): 1520–1528.

Dumont, E.R., Dávalos, L.M., Goldberg, A., Santana, S.E., Rex, K., and Voigt, C.C. (2012) Morphological innovation, diversification and invasion of a new adaptive zone. *Proceedings of the Royal Society B: Biological Sciences,* 279(1734): 1797–1805.

Duponchelle, F., Paradis, E., Ribbink, A.J., and Turner, G.F. (2008) Parallel life history evolution in mouthbrooding cichlids from the African Great Lakes. *Proceedings of the National Academy of Sciences,* 105(40): 15475–15480.

Duponchelle, F., Ribbink, A.J., Msukwa, A., Mafuka, J., Mandere, D., Bootsma, H. (2005) Food partitioning within the species-rich benthic fish community of Lake Malawi, East Africa. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(7): 1651–1664.

Durrer, L., Taborsky, M., and Frommen, J.G. (2020) Group-size preferences in a shoaling cichlid. *Behaviour*, 157(5): 415-431.

Eccles, D.H. (1974) An outline of the physical limnology of Lake Malawi (Lake Nyasa). *Limnology and Oceanography,* 19(5): 730–742.

Edgley, D.E., and Genner, M.J. (2019) Adaptive diversification of the lateral line system during cichlid fish radiation. *iScience* 16: 1–11.

Ehrlich, H. (2015) Fish skin: From clothing to tissue engineering, In Ehrlich, H. (Ed.) *Biological Materials of Marine Origin*. Biologically-Inspired Systems, Springer, Dordrecht.

El Taher, A., Böhne, A., Boileau, N., Ronco, F., Indermaur, A., Widmer, L., and Salzburger, W. (2021) Gene expression dynamics during rapid organismal diversification in African cichlid fishes. *Nature Ecology & Evolution,* 5(2): 243–250.

Engelmann, J., Hanke, W., and Bleckmann, H. (2002) Lateral line reception in still- and running water. *Journal of Comparative Physiology A,* 188(7): 513–526.

Engelmann, J., Hanke, W., Mogdans, J., and Bleckmann, H. (2000) Hydrodynamic stimuli and the fish lateral line. *Nature,* 408(6808): 51–52.

England, S.J., and Robert, D. (2022) The ecology of electricity and electroreception. *Biological Reviews,* 97(1): 383–413.

Erwin, D.H. (2015) Novelty and innovation in the history of life. *Current Biology,* 25(19): 930–940.

Fan, S., Elmer, K.R., and Meyer, A. (2012) Genomics of adaptation and speciation in cichlid fishes: recent advances and analyses in African and Neotropical lineages. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367(1587): 385.

Fan, Z., Chen, J., Zou, J., Bullen, D., Liu, C., and Delcomyn, F. (2002) Design and fabrication of artificial lateral line flow sensors. *Journal of Micromechanics and Microengineering,* 12(5): 655–661.

Faucher, K., Parmentier, E., Becco, C., Vandewalle, N., and Vandewalle, P. (2010) Fish lateral system is required for accurate control of shoaling behaviour. *Animal Behaviour,* 79: 679–687.

Feilich, K.L. (2016) Correlated evolution of body and fin morphology in the cichlid fishes. *Evolution,* 70(10): 2247–2267.

Fischer, E.K., Soares, D., Archer, K.R., Ghalambor, C.K., and Hoke, K.L. (2013) Genetically and environmentally mediated divergence in lateral line morphology in the Trinidadian guppy (*Poecilia reticulata*). *The Journal of Experimental Biology,* 216(16): 3132–3142.

Fischer, S., Oberhummer, E., Cunha-Saraiva, F., Gerber, N., and Taborsky, B. (2017) Smell or vision? the use of different sensory modalities in predator discrimination. *Behavioral Ecology and Sociobiology,* 71(10): 143.

Fitzpatrick, B.M., Fordyce, J.A., and Gavrilets, S. (2008) What, if anything, is sympatric speciation? *Journal of Evolutionary Biology,* 21(6): 1452–1459.

Fitzpatrick, B.M., Fordyce, J.A., and Gavrilets, S. (2009) Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology,* 22(11): 2342–2347.

Flores, A., Wiff, R., Ganias, K., and Marshall, C.T. (2019) Accuracy of gonadosomatic index in maturity classification and estimation of maturity ogive. *Fisheries Research,* 210(2): 50–62.

Ford, A.G.P., Rüber, L., Newton, J., Dasmahapatra, K.K., Balarin, J.D., Bruun, K., and Day, J.J. (2016) Niche divergence facilitated by fine-scale ecological partitioning in a recent cichlid fish adaptive radiation. *Evolution,* 70(12): 2718–2735.

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., Nielsen, A., and Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization methods and software,* 27(2): 233-249.

Franklin, O.D., and Morrissey, M.B. (2017) Inference of selection gradients using performance measures as fitness proxies. *Methods in Ecology and Evolution,* 8(6): 663–677.

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(1931): 20201053.

Fritzsch, B., and Straka, H. (2014) Evolution of vertebrate mechanosensory hair cells and inner ears: toward identifying stimuli that select mutation driven altered morphologies. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology,* 200(1): 5–18.

Furusawa, C., and Irie, N. (2020) Toward understanding of evolutionary constraints: experimental and theoretical approaches'. *Biophysical Reviews,* 12(5): 1155–1161.

Galis, F., and Barel, C.D.N. (1979) Comparative functional morphology of the gills of African lacustrine cichlidae (Pisces, Teleostei). *Netherlands Journal of Zoology,* 30(2): 392–430.

Galis, F., Drucker, E.G., 1996. Pharyngeal biting mechanics in centrarchid and cichlid fishes: insights into a key evolutionary innovation. *Journal of Evolutionary Biology,* 9(5): 641–670.

Gante, H.F., and Salzburger, W. (2012) Evolution: cichlid models on the runaway to speciation. *Current Biology,* 22(22): 956-958.

Gavrilets, S. (2003) Perspective: models of speciation: what have we learned in 40 years? *Evolution; International Journal of Organic Evolution,* 57(10): 2197–2215.

Gelman, S., Ayali, A., Tytell, E.D., and Cohen, A.H. (2007) Larval lampreys possess a functional lateral line system'. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology,* 193(2): 271–277.

Genner, M., Turner, G., 2011. Ancient Hybridization and Phenotypic Novelty within Lake Malawi's Cichlid Fish Radiation. *Molecular Biology and Evolution,* 29(1): 195–206.

Genner, M.J., and Turner, G.F. (2005) The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. *Fish and Fisheries,* 6(1): 1–34.

Genner, M.J., Nichols, P., Carvalho, G.R., Robinson, R.L., Shaw, P.W., Smith, A., Turner, G.F., 2007. Evolution of a cichlid fish in a Lake Malawi satellite lake. *Proceedings of the Royal Society B: Biological Sciences*, 274(1623): 2249–2257.

Genner, M.J., Seehausen, O., Lunt, D.H., Joyce, D.A., Shaw, P.W., Carvalho, G.R., and Turner, G.F. (2007) Age of cichlids: new dates for ancient lake fish radiations. *Molecular Biology and Evolution,* 24(5): 1269–1282.

Ghods, S., Waddell, S., Weller, E., Renteria, C., Jiang, H.Y., Janak, J.M., Mao, S.S., Linley, T.J., and Arola, D (2020) On the regeneration of fish scales: structure and mechanical behavior. *The Journal of Experimental Biology,* 223(10): jeb211144.

Gillespie, R.G., Bennett, G.M., De Meester, L., Feder, J.L., Fleischer, R.C., Harmon, L.J. and Hendry, A.P. (2020) Comparing adaptive radiations across space, time, and taxa. *Journal of Heredity,* 111(1): 1–20.

Ginnaw, G.M., Davidson, I.K., Harding, H.R., Simpson, S.D., Roberts, N.W., Radford, A.N., and Ioannou, C.C. (2020) Effects of multiple stressors on fish shoal collective motion are independent and vary with shoaling metric. *Animal Behaviour,* 168: 7–17.

Gould, S.J., and Eldredge, N. (1977) Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology,* 3(2): 115–151.

Grant, R.B. (2003) Evolution in Darwin's finches: a review of a study on Isla Daphne Major in the Galápagos Archipelago. *Zoology,* 106(4): 255–259.

Greenwood, A.K., Mills, M.G., Wark, A.R., Archambeault, S.L., and Peichel, C.L. (2016) Evolution of schooling behavior in threespine sticklebacks is shaped by the eda gene. *Genetics,* 203(2): 677–681.

Greenwood, A.K., Wark, A.R., Yoshida, K., and Peichel, C.L. (2013) Genetic and neural modularity underlie the evolution of schooling behavior in threespine Sticklebacks. *Current Biology*, 23: 1884-1888.

Greenwood, P.H., and Kornfield, I. (1983) The haplochromine fishes of the East African lakes: Collected papers on their taxonomy, biology and evolution (with an introduction and species index). *The Quarterly Review of Biology,* 58: 86-86.

Gregory, J.E., Iggo, A., McIntyre, A.K., and Proske, U. (1987) Electroreceptors in the platypus. *Nature,* 326(6111): 386–387.

Gregory, J.E., Iggo, A., McIntyre, A.K., and Proske, U. (1989) Responses of electroreceptors in the snout of the echidna. *The Journal of Physiology,* 414: 521–538.

Gregson, J.N.S., and de Perera, T.B. (2007) Shoaling in eyed and blind morphs of the characin Astyanax fasciatus under light and dark conditions. *Journal of Fish Biology,* 70(5): 1615–1619.

Gunter, H.M., Fan, X., Franchini, P., Fruciano, C., and Meyer, A. (2013) Shaping development through mechanical strain: the transcriptional basis of diet-induced phenotypic plasticity in a cichlid fish. *Molecular Ecology,* 22(17): 4516–4531.

Guthrie, D.M. (1986) Role of vision in fish behaviour. In Pitcher, T.J. (Ed.) *The Behaviour of Teleost Fishes*, Springer, Boston, US. 75–113.

Habel, K., Grasman, R., Gramacy, R.B., Mozharovskyi, P., and Sterratt, D.C. (2019) geometry: Mesh generation and surface tessellation. 2019. R package version 0.4.5. https://CRAN.R-project.org/package=geometry

Hara, T.J. (1986) Role of olfaction in fish behaviour. In Pitcher, T.J. (Ed.) *The Behaviour of Teleost Fishes*, Springer, Boston, US. 152–76.

Hemmings, C.C. (1966) Olfaction and vision in fish schooling. *Journal of Experimental Biology,* 45: 449–464.

Hendry, A.P., Huber, S.K., De León, L.F., Herrel, A., and Podos, J. (2009) Disruptive selection in a bimodal population of Darwin's Finches. *Proceedings of the Royal Society B: Biological Sciences,* 276(1657): 753–759.

Hensel, K., and Balon, E.K. (2001) The sensory canal systems of the living coelacanth, latimeria chalumnae: a new instalment. *Environmental Biology of Fishes,* 61(2): 117–124.

Herbert-Read, J.E., Rosén, E., Szorkhovszky, A., Ioannou, C.C., Rogell, B., Perna, A., Ramnarine, I.W., et al. (2017) How predation shapes the social interaction rules of shoaling fish. *Proceedings of the Royal Society B: Biological Sciences,* 284(1861): 20171126.

Hernández-Hernández, T., Miller, E.C., Román-Palacios, C., and Wiens, J.J. (2021) Speciation across the tree of life. *Biological Reviews,* 96(4): 1205–1242.

Higham, T., Stewart, W., and Wainwright, P. (2015) turbulence, temperature, and turbidity: the ecomechanics of predator–prey interactions in fishes. *Integrative and Comparative Biology,* 55

Hoekstra, D., and Janssen, J. (1986) Lateral line receptivity in the mottled sculpin (*Cottus bairdi*). *Ichthyology and Herpetology,* 1: 91-96.

Hoekstra, H., Hoekstra, J., Berrigan, D., Vignieri, S., Hoang, A., Hill, C., Beerli, P., and Kingsolver, J. (2001) Strength and tempo of natural selection in the wild'. *Proceedings of the National Academy of Sciences of the USA,* 98(8): 9157–9160.

Hofmann, C.M., O'Quin, K.E., Marshall,. N.J., Cronin, T.W., Seehausen, O., and Carleton, K.L. (2009) The eyes have it: regulatory and structural changes both underlie cichlid visual pigment diversity. *PLOS Biology,* 7(12): e1000266.

Holzman, R., and Hulsey, C.D. (2017) Mechanical transgressive segregation and the rapid origin of trophic novelty. *Scientific Reports,* 7(1): 40306.

Hu, Q., Wei, C., Liu, Y., and Zhao, Z. (2019) A Review of Biomimetic Artificial Lateral Line Detection Technology for Unmanned Underwater Vehicles. In Yu, H., Liu, J., Liu, L., Ju, Z., Liu, Y., and Zhou, D. (Eds.) *Intelligent Robotics and Applications*. ICIRA 2019. Lecture Notes in Computer Science, vol 11741. Springer, Cham.

Huizinga, M., Ghalambor, C.K., and Reznick, D.N. (2009) The genetic and environmental basis of adaptive differences in shoaling behaviour among populations of Trinidadian guppies, *Poecilia reticulata*. *Journal of Evolutionary Biology*, 22(9): 1860-1866.

Hulsey, C. D., and García De León, F.J. (2005) Cichlid jaw mechanics: linking morphology to feeding specialization. *Functional Ecology,* 19(3): 487–494.

Hulsey, C.D., Meyer, A., and Streelman, J.T. (2020) Convergent evolution of cichlid fish pharyngeal jaw dentitions in mollusk-crushing predators: comparative x-ray computed tomography of tooth sizes, numbers, and replacement. *Integrative and Comparative Biology,* 60(3): 656–664.

Hunter, J.P. (1998) Key innovations and the ecology of macroevolution. *Trends in Ecology & Evolution,* 13(1): 31–36.

Husemann, M., Tobler, M., McCauley, C., Ding, B., and Danley, P.D. (2017) Body shape differences in a pair of closely related Malawi cichlids and their hybrids: effects of genetic variation, phenotypic plasticity, and transgressive segregation. *Ecology and Evolution,* 7(12): 4336–4346.

Huth, A., and Wissel, C. (1994) The simulation of fish schools in comparison with experimental data. *Ecological Modelling*, State-of-the-Art in Ecological Modelling proceedings of ISEM's 8th International Conference, 75–76(9): 135–46.

Ilieş, I., Traniello, I.M., Sîrbulescu, R.F., and Zupanc, G.K.H. (2014) Determination of relative age using growth increments of scales as a minimally invasive method in the tropical freshwater *Apteronotus leptorhynchus*. *Journal of Fish Biology,* 84(5): 1312–1325.

Ioannou, C.C. (2017) Swarm intelligence in fish? The difficulty in demonstrating distributed and self-organised collective intelligence in (some) animal groups. *Behavioural Processes,* 141: 141–151.

Ioannou, C.C., Rocque, F., Herbert-Read, J.E., Duffield, C., and Firth, J.A. (2019) Predators attacking virtual prey reveal the costs and benefits of leadership. *Proceedings of the National Academy of Sciences*, 116(18): 8925-8930.

Irschick, D.J. (2002) Evolutionary approaches for studying functional morphology: examples from studies of performance capacity. *Integrative and Comparative Biology,* 42(2): 278–290.

Janssen, J. (1996) Use of the lateral line and tactile senses in feeding in four Antarctic nototheniid fishes'. *Environmental Biology of Fishes,* 47(1): 51–64.

Janssen, J. (2000) Toxicity of CO_2+: implications for lateral line studies. *Journal of Comparative Physiology A,* 186: 957–960.

Janssen, J., and Corcoran, J. (1998) Distance determination via the lateral line in the mottled sculpin. *Copeia,* 3: 657–662.

Janssen, J., Jones, W.R., Whang, A., and Oshel, P.E. (1995) Use of the lateral line in particulate feeding in the dark by juvenile alewife (*Alosa pseudoharengus*). *Canadian Journal of Fisheries and Aquatic Sciences,* 52: 358–363.

Ji, M., Zhang, Y., Zheng, X., Liu, G., and Qiu, J. (2017) A fish-shaped minimal prototype of lateral line system based on pressure sensing, 2017 IEEE International Conference on Mechatronics and Automation (ICMA), Takamatsu, Japan, 596-601.

Johnson, R.C., Nelson, G.W., Troyer, J.L., Lautenberger, J.A., Kessing, B.D., Winkler, C.A., and O'Brien, S.J. (2010) Accounting for multiple comparisons in a genome-wide association study (GWAS). *BMC Genomics,* 11(12): 724.

Jørgensen, J.M. (1985) On the fine structure of lateral-line canal organs of the herring (*Clupea harengus*). *Journal of the Marine Biological Association of the UK,* 65: 751-758.

Kagawa, K., and Takimoto, G. (2018) hybridization can promote adaptive radiation by means of transgressive segregation. *Ecology Letters,* 21(2): 264–274.

Kaler, A.S., and Purcell, L.C. (2019) Estimation of a significance threshold for genome-wide association studies. *BMC Genomics,* 20(1): 618.

Kalmijn, A.J. (1966) Electro-perception in sharks and rays. *Nature,* 212(5067): 1232–1233.

Kasumyan, A.O., and Marusov, E.A. (2007) Chemoreception in chronically anosmiated fish: a phenomenon of compensatory development of the gustatory system. *Journal of Ichthyology,* 47(8): 647–655.

Kauffman, S., and Levin, S. (1987) Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology,* 128(1): 11–45.

Kautt, A.F., Kratchowil, C.F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., Torres-Dowdall, J. et al. (2020) Contrasting signatures of genomic divergence during sympatric speciation. *Nature,* 588(7836): 106–111.

Kermen, F., Franco, L., Wyatt, C., and Yaksi, E. (2013) Neural circuits mediating olfactory-driven behavior in fish. *Frontiers in Neural Circuits,* 7(62): 1-9.

Kerr, J.R., Manes, C., and Kemp, P.S. (2016) Assessing Hydrodynamic space use of brown trout, Salmo trutta, in a complex flow environment: a return to first principles. *Journal of Experimental Biology,* 219(21): 3480–3491.

Kingsford, M.J., and Atkinson, M.H. (1994) Increments in otoliths and scales: how they relate to the age and early development of reared and wild larval and juvenile *Pagrus auratus* (Sparidae). *Marine and Freshwater Research,* 45(6): 1007–1021.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gilbert, P., and Beerli, P. (2001) The strength of phenotypic selection in natural populations. *The American Naturalist,* 157(3): 245–261.

Klein, R.J. (2007) Power analysis for genome-wide association studies. *BMC Genetics*, 8(58)

Klein, A., and Bleckmann, H. (2015) Function of lateral line canal morphology. *Integrative Zoology,* 10(1): 111–121.

Klein, A., Münz, H., and Bleckmann, H. (2013) The functional significance of lateral line canal morphology on the trunk of the marine teleost *Xiphister atropurpureus* (Stichaeidae). *Journal of Comparative Physiology A,* 199(9): 735–749.

Kmentová, N., Koblmüller, S., van Steenberge, M., Raeymaekers, J.A.M., Artois, T., De Keyzer, E.L.R., Milec, L. et al. (2020) Weak population structure and recent demographic expansion of the monogenean parasite *Kapentagyrus* spp. infecting Clupeid fishes of Lake Tanganyika, East Africa. *International Journal for Parasitology,* 50(6): 471–486.

Konings, A., 2016. *Malawi Cichlids in their Natural Habitat, 5th Edition*. Hollywood Import & Export Inc.

Konings, A., Wisor, J.M., and Stauffer Jr., J.R. (2021) Microcomputed tomography used to link head morphology and observed feeding behavior in cichlids of Lake Malaŵi. *Ecology and Evolution*, 11(9): 4605–4615.

Kottapalli, A.G.P., Asadnia, M., Miao, J., and Triantafyllou, M. (2014) Touch at a distance sensing: lateral-line inspired MEMS flow sensors. *Bioinspiration & Biomimetics*, 9(4)

Kowalko, J.E., Rohner, N., Rompani, S.B., and Peterson, B.K. (2013) Loss of schooling behaviour in cavefish through sight-dependent and sight-independent mechanisms. *Current Biology*, 23: 1874-1883.

Kowalko, J.E., Rohner, N., Rompani, S.B., Peterson, B.K., Linden, T.A., Yoshizawa, M., Kay, E.M., et al. (2013) Loss of schooling behavior in cavefish through sight-dependent and sight-independent mechanisms. *Current Biology,* 23(19): 1874–1883.

Kratochwil, C.F., Liang, Y., Gerwin, J., Woltering, J.M., Urban, S., Henning, F., Machado-Schiaffino, G., Hulsey, C.D., and Meyer, A. (2018) Agouti-related peptide 2 facilitates convergent evolution of stripe patterns across cichlid fish radiations. *Science,* 362(6413): 457–460.

Kroese A., and Schellart, N. (1992) Velocity- and acceleration-sensitive units in the Posterior lateral line of the trout. *Journal of Neurophysiology,* 68: 2212-2221.

Kulpa, M., Bak-Coleman, J., and Coombs, S. (2015) The lateral line is necessary for blind cavefish rheotaxis in non-uniform flow. *Journal of Experimental Biology,* 218(10): 1603–1612.

Ladich, F., and Schulz-Mirbach, T. (2016) Diversity in fish auditory systems: one of the riddles of sensory biology. *Frontiers in Ecology and Evolution,* 4(28): 1–26

Lai, Y., Yeung, C.K.L., Omland, K.E., Pang, E., Hao, Y., Liao, B., Cao, H. et al. (2019) Standing genetic variation as the predominant source for adaptation of a songbird. *Proceedings of the National Academy of Sciences,* 116(6): 2152–2157.

Lande, R. (1982) Rapid origin of sexual isolation and character divergence in a cline. *Evolution,* 36(2): 213–223.

Lande, R., and Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution,* 37(6): 1210–1226.

Lannoo, M.J. (1988) The evolution of the amphibian lateral line system and its bearing on amphibian phylogeny. *Journal of Zoological Systematics and Evolutionary Research,* 26(2): 128–134.

Laurà, R., Abbate, F., Germanà, G.P., Montalbano, G., Germanà, A., and Levanti, M. (2018) Fine structure of the canal neuromasts of the lateral line system in the adult zebrafish. *Anatomia, Histologia, Embryologia,* 47(4): 322–329.

Li, C., Wang, X., Wu, J., Zhang, X., Fan, C., Guo, H., and Song, J. (2018) Heterogeneity of neuromasts in a fish without lateral line canals: the pufferfish (Takifugu obscurus) model. *Journal of Experimental Biology,* 221(19): jeb186163.

Li, G., Kolomenskiy, D., Liu, H., Thiria, B., and Godoy-Diana, R. (2022) Hydrodynamical fingerprint of a neighbour in a fish lateral line. *Frontiers in Robotics and AI*, 9.

Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics,* 27(21): 2987–2993.

Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. https://doi.org/10.48550/arXiv.1303.3997

Li, H., and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25(14): 1754-1760.

Liao, J.C. (2006) The role of the lateral line and vision on body kinematics and hydrodynamic preference of rainbow trout in turbulent flow. *The Journal of Experimental Biology,* 209(20): 4077–4090.

Liao, J.C. 2007. A review of fish swimming mechanics and behaviour in altered flows. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362: 1973–1993.

Linjia, J., Romero-Carvajal, A., Haug, J.S., Seidel, C.W., and Piotrowski, T. (2014) Gene-expression analysis of hair cell regeneration in the zebrafish lateral line. *Proceedings of the National Academy of Sciences of the USA,* 111(14): 1383–1392.

Lippitsch, E. (1990) Scale morphology and squamation patterns in cichlids (Teleostei, Perciformes): a comparative study. *Journal of Fish Biology,* 37(2): 265–291.

Liu, G., Wang, A., Wang, X., and Liu, P. (2016) A review of artificial lateral line in sensor fabrication and bionic applications for robot fish. *Applied Bionics and Biomechanics,* 2016: 4732703.

Liu, Z., Chen, T., Zhu, T., Zeng, Z., Lyu, Z., Wang, J., Messenger, K. et al. (2018) Prevalence of cryptic species in morphologically uniform taxa – fast speciation and evolutionary radiation in Asian frogs. *Molecular Phylogenetics and Evolution,* 127(10): 723–731.

Lloyd, E., Olive, C., Stahl, B.A., Jaggard, J.B., Amaral, P., Duboué, E.R., and Keene, A.C. (2018) Evolutionary shift towards lateral line dependent prey capture behavior in the blind Mexican cavefish. *Developmental Biology,* 441(2): 328–337.

Lutz, V., Stratz, P., Preuß, S., Tetens, J., Grashorn, M.A., Bessei, W., and Bennewitz, J. (2017) A genome-wide association study in a large F2-cross of laying hens reveals novel genomic regions associated with feather pecking and aggressive pecking behavior. *Genetics, Selection, Evolution,* 49(1): 18.

Maan, M.E., and Sefc, K.M. (2013) Colour variation in cichlid fish: developmental mechanisms, selective pressures and evolutionary consequences. *Seminars in Cell & Developmental Biology,* 24(6–7): 516–528.

Maan, M.E., Seehausen, O., and Groothuis, T.G.G. (2017) Differential survival between visual environments supports a role of divergent sensory drive in cichlid fish speciation. *The American Naturalist,* 189(1): 78–85.

Maan, M.E., Seehausen, O., Söderberg, L., Johnson, L., Ripmeester, E.A.P., Mrosso, H.D.J., Taylor, M.I., van Dooren, T.J.M and van Alphen, J.J.M. (2004) Intraspecific sexual selection on a speciation trait, male coloration, in the Lake Victoria cichlid *Pundamilia nyererei*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1556): 2445–2452.

Maeda, K., Takeda, M., Kamiya, K., Aibara, M., Mzighani, S.I., Nishida, M., Mizoiri, S. et al. (2009) Population structure of two closely related pelagic cichlids in Lake Victoria, *Haplochromis pyrrhocephalus* and *H. laparogramma*. *Gene,* 441(1–2): 67–73.

Magalhaes, I.S., D'Agostino, D., Hohenlohe, P.A., and MacColl, A.D.C. The ecology of an adaptive radiation of three-spined stickleback from North Uist, Scotland. *Molecular Ecology,* 25(17): 4319–4336.

Magurran, A.E. (1990) The adaptive significance of schooling as an anti-predator defence in fish. *Annales Zoologici Fennici*, 27: 51–66.

Malinsky, M., Matschiner, M., and Svardal, H. (2021) Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources,* 21(2): 584–595.

Malinsky, M., and Salzburger, W. (2016) Environmental context for understanding the iconic adaptive radiation of cichlid fishes in Lake Malawi. *Proceedings of the National Academy of Sciences of the USA,* 113: 11654-11656.

Malinsky, M., Challis, R.J., Tyers, A.M., Schiffels, S., Terai, Y., Ngatunga, B.P., Miska, E.A., Durbin, R., Genner, M.J. and Turner, G.F. (2015) Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science,* 350(6267): 1493–1498.

Malinsky, M., Svardal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F., and Durbin, R. (2018) Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecology & Evolution,* 2(12): 1940–1955.

Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W. (2010) Robust relationship inference in genome-wide association studies. *Bioinformatics,* 26(22): 2867–2873.

Marees, A.T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., and Derks, E.M. (2018) A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *International Journal of Methods in Psychiatric Research,* 27(2).

Marlin, T., Snekser, J.L., and Leese, J.M. (2019) Juvenile convict cichlids shoaling decisions in relation to shoal size and age. *Ethology*, 125: 485– 491.

Marques, D.A., Jones, F.C., Di Palma, F., Kingsley, D.M., and Reimchen, T.E. (2018) experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nature Ecology & Evolution,* 2(7): 1128–1138.

Marranzino, A.N., and Webb, J.F. (2018) Flow sensing in the deep sea: the lateral line system of Stomiiform fishes. *Zoological Journal of the Linnean Society,* 183(4): 945–965.

Marshall, N.J., 1996. The lateral line systems of three deep-sea fish. *Journal of Fish Biology,* 49: 239–258.

Martin, C.H. (2012) Weak disruptive selection and incomplete phenotypic divergence in two classic examples of sympatric speciation: Cameroon crater lake cichlids. *The American Naturalist,* 180(4): 90–109.

Martin, C.H., and Gould, K.J. (2020) Surprising spatiotemporal stability of a multi-peak fitness landscape revealed by independent field experiments measuring hybrid fitness. *Evolution Letters,* 4(6): 530–544.

Martin, C.H., and Wainwright, P.C. (2013) Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science,* 339(6116): 208–211.

Martin, R.A., and Pfennig, D.W. (2009) Disruptive selection in natural populations: the roles of ecological specialization and resource competition. *The American Naturalist,* 174(2): 268–281.

Martínez Medina, L., Garcia, C.M., Urbina, A.F., Manjarrez, J., and Moyaho, A. (2013) Female vibration discourages male courtship behaviour in the amarillo fish (*Girardinichthys multiradiatus*). *Behavioural Processes,* 100(11): 163–168.

Masonick, P., Meyer, A., and Hulsey, C.D. (2022) Phylogenomic analyses show repeated evolution of hypertrophied lips among Lake Malawi cichlid fishes. *Genome Biology and Evolution,* 14(4): evac051

Matschiner, M., Böhne, A., Ronco, F., and Salzburger, W. (2020) The genomic timeline of cichlid fish diversification across continents. *Nature Communications,* 11(5895)

McDermott, B.M., Baucom, J.M., and Hudspeth, A.J. (2007) Analysis and functional evaluation of the hair-cell transcriptome. *Proceedings of the National Academy of Sciences of the USA,* 104(28): 11820–11825.

McGee, M.D., Borstein, S.R., Meier, J.I., Marques, D.A., Mwaiko, S., Taabu, A., Kishe, M.A. et al. (2020) The ecological and genomic basis of explosive adaptive radiation. *Nature,* 586(7827): 75–79.

McHenry, M.J. and van Netten, S.M. (2007) The flexural stiffness of superficial neuromasts in the zebrafish (*Danio rerio*) lateral line. *Journal of Experimental Biology*. 210(23): 4244–4253

McHenry, M.J., and Liao, J.C. (2014) The hydrodynamics of flow stimuli. In Coombs, S., Bleckmann, H., Fay, R.R., and Popper, A.N. (Eds.) *The Lateral Line System*, 73–98. Springer Handbook of Auditory Research. Springer, New York, USA.

McKee, A., Soto, A.P., Chen, P., and McHenry, M.J. (2020) the sensory basis of schooling by intermittent swimming in the rummy-nose tetra (*Hemigrammus rhodostomus*). *Proceedings of the Royal Society B: Biological Sciences,* 287(1937): 20200568.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K. et al. (2010) The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research,* 20(9): 1297–1303.

Medina, L.M., Garcia, C.M., Urbina, A.F., Manjarrez, J., and Moyaho, A. (2013) Female vibration discourages male courtship behaviour in the Amarillo fish (*Girardinichthys multiradiatus*). *Behavioural Processes,* 100: 163-168.

Mehta, T.K., Koch, C., Nash, W., Knaack, S.A., Sudhakar, P., Olbei, M., Bastkowski, S., Penso-Dolfin, L., Korcsmaros, T., Haerty, W., Roy, S., Di-Palma, F. (2021) Evolution of regulatory networks associated with traits under selection in cichlids. *Genome Biology,* 22(25)

Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L., and Seehausen, O. (2017) Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications,* 8(1): 14363.

Meier, J.I., Marques, D.A., Wagner, C.E., Excoffier, L., and Seehausen, O. (2018) Genomics of parallel ecological speciation in Lake Victoria cichlids. *Molecular Biology and Evolution,* 35(6): 1489–1506.

Mekdara, P.J., Schwalbe, M.A.B., Coughlin, L.L., and Tytell, E.D. (2018) the effects of lateral line ablation and regeneration in schooling giant danios. *Journal of Experimental Biology,* 221(8): jeb175166.

Mekdara, P.J., Tirmizi, S., Schwalbe, M.A.B., and Tytell, E.D. (2022) Comparison of aminoglycoside antibiotics and cobalt chloride for ablation of the lateral line system in giant danios. *Integrative Organismal Biology,* 4(1): obac012.

Michel, A.P., Sim, S., Powell, T.H.Q., Taylor, M.S., Nosil, P., and Feder, J.L. (2010) Widespread genomic divergence during sympatric speciation'. *Proceedings of the National Academy of Sciences,* 107(21): 9724–9729.

Miller, A.H., Stroud, J.T., and Losos, J.B. (2022) The ecology and evolution of key innovations. *Trends in Ecology & Evolution*, https://doi.org/10.1016/j.tree.2022.09.005.

Miller, N., and Gerlai, R. (2012) Automated tracking of zebrafish shoals and the analysis of shoaling behavior. In Kalueff, A.V. and Stewart, A.M. (Eds.) *Zebrafish Protocols for Neurobehavioral Research*, Neuromethods, Humana Press, Totowa, NJ. 217–230.

Mills, M.G., Greenwood, A.K., and Peichel, C.L. (2014) Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo,* 5(1): 5

Mirjany, M., Preuss, T., and Faber, D.S. (2011) Role of the lateral line mechanosensory system in directionality of goldfish auditory evoked escape response. *The Journal of Experimental Biology,* 214(20): 3358–3367.

Miyagi, R., Terai, Y., Aibara, M., Sugawara, T., Imai, H., Hidenori, T., Mzighani, S.I., Okitsu, T., Wada, A., and Okada, N. (2012) Correlation between nuptial colors and visual sensitivities tuned by opsins leads to species richness in sympatric Lake Victoria cichlid fishes. *Molecular Biology and Evolution,* 29(11): 3281–3296.

Mo, W., and Nicolson, T. (2011) Both pre- and postsynaptic activity of Nsf prevents degeneration of hair-cell synapses. *PLOS One,* 6(11): e27146.

Mogdans, J. (2019) Sensory ecology of the fish lateral-line system: morphological and physiological adaptations for the perception of hydrodynamic stimuli'. *Journal of Fish Biology,* 95(1): 53–72.

Mogdans, J., and Bleckmann, H. (2012) Coping with flow: behavior, neurophysiology and modeling of the fish lateral line system. *Biological Cybernetics,* 106(11–12): 627–642.

Montgomery, J., 1989. Lateral Line Detection of Planktonic Prey. In Coombs, S., Görner, P., Münz, H. (Eds.), *The Mechanosensory Lateral Line*. Springer, New York, NY, 561-574

Montgomery, J.C., and Clements, K. (2000) Disaptation and recovery in the evolution of Antarctic fishes. *Trends in Ecology & Evolution,* 15(7): 267–271.

Montgomery, J.C., and Coombs, S. (1998) Peripheral encoding of moving sources by the lateral line system of a sit-and-wait predator. *Journal of Experimental Biology,* 201(1): 91–102.

Montgomery, J.C., and MacDonald, J.A. (1987) Sensory tuning of lateral line receptors in Antarctic fish to the movements of planktonic prey. *Science,* 235(4785): 195–196.

Montgomery, J.C., Baker, C.F. and Carton, A.G. (1997) The lateral line can mediate rheotaxis in fish. *Nature* 389(10): 960–963.

Montgomery, J.C., Bleckmann, H., and Coombs, S. (2014) Sensory ecology and neuroethology of the lateral line. In Coombs, S., Bleckmann, H., Fay, R.R., and Popper, A.N. (Eds.) *The Lateral Line System*, Springer Handbook of Auditory Research. Springer, New York, USA. 121–150.

Montgomery, J.C., Coombs, S. and Halstead, M. (1995) Biology of the mechanosensory lateral line in fishes. *Reviews in Fish Biology and Fisheries,* 5(4): 399–416.

Montgomery, J.C., Coombs, S., and Janssen, J. (1994) Form and function relationships in lateral line systems: comparative data from six species of Antarctic notothenioid fish. *Brain, Behavior and Evolution,* 44(6): 299–306.

Moorad, J.A., and Wade, M.J. (2013) Selection gradients, the opportunity for selection, and the coefficient of determination. *The American Naturalist,* 181(3): 291–300.

Moser, F.N., van Rijssel, J.C., Mwaiko, S., Meier, J.I., Ngatunga, B., and Seehausen, O. (2018) The onset of ecological diversification 50 years after colonization of a crater lake by Haplochromine cichlid fishes. *Proceedings of the Royal Society B: Biological Sciences,* 285(1884): 20180171.

Moss, S., Tittaferrante, S., Way, G.P., Fuller, A., Sullivan, N., Ruhl, N., and McRobert, S.P. (2015) Interactions between aggression, boldness and shoaling within a brood of convict cichlids (*Amatitlania nigrofasciatus*). *Behavioural Processes*, 121: 63-69.

Müller, U., and Schwartz, E. (1982) Influence of single neuromasts on prey localizing behavior of the surface feeding fish, *Aplocheilus lineatus*. *Journal of Comparative Physiology* 149(3): 399–408.

Munby, H., Linderoth, T., Fischer, B., Du, M., Vernaz, G., Tyers, A.M., Ngatunga, B.P. et al. (2021) Differential use of multiple genetic sex determination systems in divergent ecomorphs of an African crater lake cichlid. *bioRxiv*. https://doi.org/10.1101/2021.08.05.455235.

Muschick, M., Barluenga, M., Salzburger, W., and Meyer, A. (2011) Adaptive phenotypic plasticity in the midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. *BMC Evolutionary Biology,* 11(1): 116.

Muschick, M., Indermaur, A., and Salzburger, W. (2012) convergent evolution within an adaptive radiation of cichlid fishes. *Current Biology,* 22(24): 2362–2368.

Muschick, M., Nosil, P., Roesti, M., Dittmann, M.T., Harmon, L., and Salzburger, W. (2014) Testing the stages model in the adaptive radiation of cichlid fishes in East African Lake Tanganyika. *Proceedings of the Royal Society B: Biological Sciences,* 281(1795): 20140605.

Musilová, Z., Indermaur, A., Bitja-Nyom, A.R., Omelchenko, D., Kłodawska, M., Albergati, L., Remišová, K., Salzburger, W., 2019. Evolution of the visual sensory system in cichlid fishes from crater lake Barombi Mbo in Cameroon. *Molecular Ecology,* 28(23): 5010–5031.

Nakae, M., and Sasaki, K. (2010) lateral line system and its innervation in tetraodontiformes with outgroup comparisons: descriptions and phylogenetic implications. *Journal of Morphology,* 271(5): 559–579.

Navon, D., Hatini, P., Zogbaum, L., Albertson, R.C., 2021. The genetic basis of coordinated plasticity across functional units in a Lake Malawi cichlid mapping population. *Evolution*, 75(3): 672–687.

Navon, D., Olearczyk, N., and Albertson, R.C. (2017) Genetic and developmental basis for fin shape variation in African cichlid fishes. *Molecular Ecology,* 26(1): 291–303.

New, J.G., Fewkes, L.A., and Khan, A.N. (2001) Strike feeding behavior in the muskellunge, *Esox masquinongy*: contributions of the lateral line and visual sensory systems. *Journal of Experimental Biology*, 204(6): 1207-1221.

Nickles, K.R., Hu, Y., Majoris, J.E., Buston, P.M., Webb, J.F., 2020. Organization and ontogeny of a complex lateral line system in a goby (*Elacatinus lori*), with a Consideration of Function and Ecology. *Ichthyology & Herpetology*, 108(4): 863–885.

Nonaka, E., Svanbäck, R., Thibert-Plante, X., Englund, G., and Brännström, Å. (2015) Mechanisms by which phenotypic plasticity affects adaptive divergence and ecological speciation. *The American Naturalist,* 186(5): 126-143.

Northcutt, R.G. (1986) Lungfish neural characters and their bearing on Sarcopterygian phylogeny'. *Journal of Morphology,* 190(S1): 277–297.

Nosil, P. (2005) The role of selection and gene flow in the evolution of sexual isolation in Timema walking sticks and other Orthopteroids. *Journal of Orthoptera Research*, 14(2): 247–253

O'Quin, C.T., Drilea, A.C., Conte, M.A., and Kocher, T.D. (2013) Mapping of pigmentation QTL on an anchored genome assembly of the cichlid fish, *Metriaclima zebra. BMC Genomics,* 14(1): 287.

Obolski, U., Ram, Y., and Hadany, L. (2018) Key issues review: evolution on rugged adaptive landscapes. *Reports on Progress in Physics. Physical Society (Great Britain),* 81(1): 012602.

Oliver, M.K. (1984) Systematics of African cichlid fishes: Determination of the most primitive taxon, and studies on the haplochromines of Lake Malawi (Teleostei: Cichlidae). Unpublished Ph.D. dissertation, Yale University, New Haven, Connecticut.

Olszewski, J., Haehnel, M., Taguchi, M., and Liao, J.C. (2012) Zebrafish larvae exhibit rheotaxis and can escape a continuous suction source using their lateral line. *PLOS One,* 7(5): 36661.

Orr, H.A. (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics,* 6(2): 119–127.

Parin, N.V., and Astakhov, D.A. Studies on the acoustico-lateralis system of Beloniform fishes in connection with their systematics. Copeia, 1982(2): 276–291.

Parry, J.W.L., Carleton, K.L., Spady, T., Carboo, A., Hunt, D.M., and Bowmaker, J.K. (2005) Mix and match color vision: tuning spectral sensitivity by differential opsin gene expression in Lake Malawi cichlids. *Current Biology,* 15(19): 1734–1739.

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: 641501.

Parsons, K.J., Márquez, E., and Albertson, R.C. (2012) Constraint and opportunity: the genetic basis and evolution of modularity in the cichlid mandible. *The American Naturalist,* 179(1): 64–78.

Partridge, B.L., and Pitcher, T.J. (1980) The sensory basis of fish schools: relative roles of lateral line and vision. *Journal of Comparative Physiology,* 135(4): 315–325.

Patton, A.H., Richards, E.J., Gould, K.J., Buie, L.K., and Martin, C.H, (2022) Hybridization alters the shape of the genotypic fitness landscape, increasing access to novel fitness peaks during adaptive radiation. *eLife,* 11(5): e72905.

Pavlov, D.S., and Kasumyan, A.O. (2000) Patterns and mechanisms of schooling behavior in fish: a review. *Journal of Ichthyology*, 40(2): S163.

Peach, M.B., and Rouse, G.W. (2000) The morphology of the pit organs and lateral line canal neuromasts of *Mustelus antarcticus* (Chondrichthyes: Triakidae). *Journal of the Marine Biological Association of the United Kingdom,* 80(1): 155–162.

Pearson, K. (1902) Mathematical contributions to the theory of evolution. xi.—on the influence of natural selection on the variability and correlation of organs. *Proceedings of the Royal Society of London,* 69(451–458): 330–333.

Pérez-Escudero, A., Vicente-Page, J., Hinz, R.C., Arganda, S., and de Polavieja, G.G. (2014) idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods,* 11: 743–748.

Petren, K., Grant, B.R., and Grant, P.R. (1999) A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proceedings of the Royal Society B: Biological Sciences,* 266(1417): 321.

Pfaender, J., Hadiaty, R.K., Schliewen, U.K., and Herder, F. (2016) Rugged adaptive landscapes shape a complex, sympatric radiation. *Proceedings of the Royal Society B: Biological Sciences,* 283(1822): 20152342.

Pita, D., Moore, B.A., Tyrrell, L.P., and Fernández-Juricic, E. (2015) Vision in two cyprinid fish: implications for collective behavior. *PeerJ*, 3: e1113.

Pitcher, T. (1979) Sensory information and the organization of behaviour in a shoaling cyprinid fish. *Animal Behaviour* 27(2): 126–149.

Pitcher, T.J. (1983) Heuristic definitions of fish shoaling behaviour. *Animal Behaviour,* 31: 611–613.

Pitcher, T.J. (1986) Functions of shoaling behaviour in teleosts. In: *Behaviour of Teleost Fishes*, Pitcher, T.J., Eds. (Springer, Boston, MA), 363–427.

Pitcher, T.J., Partridge, B.L., and Wardle, C.S. (1976) A blind fish can school. *Science,* 194: 963–965.

Plenderleith, M., van Oosterhout, C., Robinson, R.L., and Turner, G.F. (2005) Female preference for conspecific males based on olfactory cues in a Lake Malawi cichlid fish. *Biology Letters,* 1(4): 411–414.

Pohlmann, K., Atema, J., and Breithaupt, T. (2004) The importance of the lateral line in nocturnal predation of piscivorous catfish. *The Journal of Experimental Biology,* 207(17): 2971–2978.

Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., van der Auwera, G.A., Kling, D.E. et al. (2018) Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*. https://doi.org/10.1101/201178

Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., van der Auwera, G.A., Kling, D.E. et al. (2018) Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*. https://doi.org/10.1101/201178.

Prats, E., Gómez-Canela, C., Ben-Lulu, S., Ziv, T., Padrós, F., Tornero, D., Garcia-Reyero, N., Tauler, R., Admon, A., and Raldúa, D. (2017) Modelling acrylamide acute neurotoxicity in zebrafish larvae. *Scientific Reports,* 7(1): 13952.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics,* 38(8): 904–909.

Price, A.L., Zaitlen, N.A., Reich, D., and Patterson, N. (2010) New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics,* 11(7): 459–463.

Quinlan, A.R., and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26(6): 841–842. https://github.com/arq5x/bedtools2

R Core Team, (2020) R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rainey, P.B., and Travisano, M. (1998) Adaptive radiation in a heterogeneous environment. *Nature,* 394(6688): 69–72.

Rajkov, J., Weber, A.A., Salzburger, W., and Egger, B. (2018) Adaptive phenotypic plasticity contributes to divergence between lake and river populations of an East African cichlid fish. *Ecology and Evolution,* 8(15): 7323–7333.

Reaney, A.M., Bouchenak-Khelladi, Y., Tobias, J.A., and Abzhanov, A. (2020) Ecological and morphological determinants of evolutionary diversification in Darwin's finches and their relatives. *Ecology and Evolution,* 10(24): 14020–14032.

Ribbink, A.J., Marsh, B.A., Marsh, A.C., Ribbink, A.C., Sharp, B.J., 1983. A preliminary survey of the cichlid fishes of rocky habitats in Lake Malawi. *South African Journal of Zoology*, 18(3): 149–310.

Rieseberg, L.H., Archer, M.A., and Wayne, R.K. (1999) Transgressive segregation, adaptation and speciation. *Heredity,* 83(4): 363–372.

Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., and Firth, D. (2022) MASS: Support Functions and Datasets for Venables and Ripley's MASS. https://CRAN.R-project.org/package=MASS.

Robertson, S., Bradley, J.E., and MacColl, A.D.C. (2017) Eda haplotypes in three-spined stickleback are associated with variation in immune gene expression. *Scientific Reports,* 7(1): 42677.

Robinson, R.L., and Turner, G.F. (1990) Comparison of the social behaviour of an endemic Lake Malawi cichlid in the field and laboratory. *Journal of Fish Biology*, 36: 615-616.

Rohlf, F.J. (2015) The tps series of software. *Hystrix, the Italian Journal of Mammalogy,* 26(1): 9–12.

Romenskyy, M., Herbert-Read, J.E., Ioannou, C.C., Szorkovszky, A., Ward, A.J., and Sumpter, D.J. (2020) Quantifying the structure and dynamics of fish shoals under predation threat in three dimensions. *Behavioral Ecology*, 31(2): 311-321.

Ronco, F., and Salzburger, W. (2021) Tracing evolutionary decoupling of oral and pharyngeal jaws in cichlid fishes. *Evolution Letters,* 5(6): 625–635.

Ronco, F., Büscher, H.H., Indermaur, A., and Salzburger, W. (2020) The taxonomic diversity of the cichlid fish fauna of ancient Lake Tanganyika, East Africa. *Journal of Great Lakes Research,* 46(5): 1067–1078.

Ronco, F., Matschiner, M., Böhne, A., Boila, A., Büscher, H.H., El Taher, A., Indermaur, A. et al. (2021) Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature,* 589(7840): 76–81.

Roussel, J., Barber, C.B., Habel, K., Grasman, R., Gramacy, R.B., Mozharovskyi, P., Sterratt, D.C. (2022) geometry: mesh generation and surface tessellation. https://cran.r-project.org/web/packages/geometry/geometry.pdf

Rowe, D.K., Dean, T.L., Williams, E., and Smith, J.P. (2003) Effects of turbidity on the ability of juvenile rainbow trout, Oncorhynchus mykiss, to feed on limnetic and benthic prey in laboratory tanks. *New Zealand Journal of Marine and Freshwater Research,* 37: 45-52.

Rundell, R.J., and Price, T.D. (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution,* 24(7): 394–399.

Rundle, H.D., and Nosil, P. (2005) Ecological speciation. *Ecology Letters,* 8(3): 336–352.

Sabbah, S., Laria, R.L., Gray, S.M., and Hawryshyn, C.W. (2010) Functional diversity in the color vision of cichlid fishes. *BMC Biology,* 8(1): 133.

Salzburger, W. (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Molecular Ecology,* 18(2): 169–185.

Salzburger, W., Mack, T., Verheyen, W., and Meyer, A. (2005) Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology,* 5(17): 237–274.

Santos, M.E., Braasch, I., Bioleau, N., Meyer, B.S., Sauteur, L., Böhne, A., Belting, H., Affolter, M., and Salzburger, W. (2014) The evolution of cichlid fish egg-spots is linked with a cis-regulatory change. *Nature Communications,* 5(1): 5149.

Satou, M., Takeuchi, H., Nishii, J., Tanabe, M., Kitamura, S., Okumotit, N., and Iwata, M. (1994) Behavioral and electrophysiological evidences that the lateral line is involved in the inter-sexual

vibrational communication of the Himé Salmon (landlocked red salmon, *Oncorhynchus nerka*). *Journal of Comparative Physiology A,* 174(5): 539–549.

Saunders, A.J., and Montgomery, J.C. (1985) Field and laboratory studies of the feeding behaviour of the piper *Hyporhamphus ihi* with reference to the role of the lateral line in feeding. *Proceedings of the Royal society of London. Series B. Biological sciences*, 224(1235): 209-221.

Savolainen, V., Anstett, M., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N., and Baker, W.J. (2006) Sympatric speciation in palms on an oceanic island. *Nature,* 441(7090): 210–213.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods,* 9, 676–682.

Schluter, D. (1993) Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology,* 74(3): 699–709.

Schluter, D. (2000) *The Ecology of Adaptive Radiation*. Oxford Series in Ecology and Evolution. Oxford University Press, Oxford, New York.

Schluter, D. (2009) Evidence for ecological speciation and its alternative. *Science,* 323(5915): 737–741.

Schluter, D., and Nagel, L.M. (1995) Parallel speciation by natural selection. *The American Naturalist,* 146(2): 292–301.

Schluter, D., and Nychka, D. (1994) Exploring fitness surfaces. *The American Naturalist,* 143(4): 597–616.

Schluter, D., and Price, T. (1993) Honesty, perception and population divergence in sexually selected traits. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* 253(1336): 117–122.

Schluter, D., Marchinko, K.B., Arnegard, M.E., Zhang, H., Brady, S.D., Jones, F.C., Bell, M.A., and Kingsley, D.M. (2021) Fitness maps to a large-effect locus in introduced stickleback populations. *Proceedings of the National Academy of Sciences,* 118(3).

Schluter, D., Rambaut, A., Clarke, B.C., and Grant, P.R. (1996) Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* 351(1341): 807–814.

Schmitz, A., Bleckmann, H., and Mogdans, J. (2014) The lateral line receptor array of cyprinids from different habitats. *Journal of Morphology,* 275(4): 357–70.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9: 671–675.

Schneider, R.F., and Meyer, A. (2017) How plasticity, genetic assimilation and cryptic genetic variation may contribute to adaptive radiations. *Molecular Ecology,* 26(1): 330–350.

Schwalbe, M.A.B, and Webb, J.F. (2015) The effect of light intensity on prey detection behavior in two Lake Malawi cichlids, *Aulonocara stuartgranti* and *Tramitichromis* sp. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology,* 201(4): 341–356.

Schwalbe, M.A.B., and Webb, J.F. (2014) Sensory basis for detection of benthic prey in two Lake Malawi cichlids. *Zoology,* 117(2): 112–121.

Schwalbe, M.A.B., Bassett, D.K., and Webb, J.F. (2012) Feeding in the dark: lateral-line-mediated prey detection in the peacock cichlid *Aulonocara stuartgranti*. *Journal of Experimental Biology,* 215(12): 2060–2071.

Schwalbe, M.A.B., Sevey, B.J., and Webb, J.F. (2016) Detection of artificial water flows by the lateral line system of a benthic feeding cichlid fish. *Journal of Experimental Biology,* 219: 1050-1059.

Seehausen, M., and van Alphen, J.J.M. (1999) Evolution of colour patterns in East African cichlid fish. *Journal of Evolutionary Biology,* 12(3): 514–534.

Seehausen, O. (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences,* 273(1597): 1987–1998.

Seehausen, O. (2007) Evolution and ecological theory: chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. *Heredity,* 99: 361–363

Seehausen, O. (2015) Process and pattern in cichlid radiations – inferences for understanding unusually high rates of evolutionary diversification. *New Phytologist,* 207(2): 304–312.

Seehausen, O., and Schluter, D. (2004) Male-male competition and nuptial-colour displacement as a diversifying force in Lake Victoria cichlid fishes. *Proceedings of the Royal Society B: Biological Sciences,* 271(1546): 1345–1353

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L. et al. (2014) Genomics and the origin of species. *Nature Reviews Genetics,* 15(3): 176–192.

Seehausen, O., Terai, Y., Magalhaes, I.S., Carleton, K.L., Mrosso, M.D.J., Miyagi, R., van der Sluijs, I. et al. (2008) Speciation through sensory drive in cichlid fish. *Nature,* 455(7213): 620–626.

Seleit, A., Ansai, S., Yamahira, K., Masengi, K.W.A., Naruse, K., and Centanin, L. (2021) Diversity of lateral line patterns and neuromast numbers in the genus *Oryzias*. *Journal of Experimental Biology*, 224(24)

Selz, O.M., Pierotti, M.E.R., Maan, M.E., Schmid, C., and Seehausen, O. (2014) Female preference for male color is necessary and sufficient for assortative mating in two cichlid sister species. *Behavioral Ecology* 25(3): 612–626.

Selz, O.M., Thommen, R., Pierotti, M.E.R., Anaya-Rojas, J.M., and Seehausen O. (2016) Differences in male coloration are predicted by divergent sexual selection between populations of a cichlid fish. *Proceedings of the Royal Society B: Biological Sciences,* 283(1830): 20160172.

Sharov, A.A. (2014) Evolutionary constraints or opportunities? *BioSystems*, 4: 21–30.
Shaw, E. (1962) The schooling of fishes. *Scientific American,* 206: 128–141.

Shoaib, M., Hamid, N.H., Malik, A.F., Ali, N.B.Z., and Jan, M.T. (2016) A review on key issues and challenges in devices level MEMS testing. *Journal of Sensors*, 1639805: 1–14.

Siby, P., Machado, J.P., Maldonado, E., Vasconcelos, V., O'Brien, S.J., Johnson, W.E., and Antunes, A. (2012) Fish lateral line innovation: insights into the evolutionary genomic dynamics of a unique mechanosensory organ. *Molecular Biology and Evolution,* 29(12): 3887–3898.

Simmons, A.M., Costa, L.M., and Gerstein, H.B. (2004) Lateral line-mediated rheotactic behavior in tadpoles of the African clawed frog (Xenopus laevis). *Journal of Comparative Physiology A,* 190(9): 747-758.

Simpson, G.G. (1953) *The Major Features of Evolution*. Columbia University Press.

Skandalis, D.A., Lunsford, E.T., and Liao, J.C. (2021) Corollary discharge enables proprioception from lateral line sensory feedback. *PLOS Biology,* 19(10): e3001420.

Smith, J.M. (1966) Sympatric Speciation. *The American Naturalist,* 100(916): 637–650.

Sobel, J.M., Chen, G.F., Watt, L.R., and Schemske, D.W. (2010) The biology of speciation. *Evolution,* 64(2): 295–315.

Song, J., Yan, H.Y., and Popper, A.N. (1995) Damage and recovery of hair cells in fish canal (but not superficial) neuromasts after gentamicin exposure. *Hearing Research,* 91: 63-71.

Sparks, J.S. (2008) Phylogeny of the cichlid subfamily Etroplinae and taxonomic revision of the Malagasy cichlid genus *Paretroplus* (Teleostei: Cichlidae). *Bulletin of the American Museum of Natural History,* 314(1): 1–151

Stephens, R.R. (1985) The lateral line system of the gizzard shad, *Dorosoma cepedianum* Lesueur (Pisces: Clupeidae). *Copeia,* 540-556.

Stewart, L. (2019) Open Foam Tutorial: Simulation with 3D Geometry (.stl), https://www.youtube.com/watch?v=oc4V79F78AQ

Stewart, W.J., Johansen, J.L., and Liao, J.C. (2017) A non-toxic dose of cobalt chloride blocks hair cells of the zebrafish lateral line. *Hearing Research,* 350: 17-21.

Stewart, W.J., Nair, A., Jiang, H., and McHenry, M. (2014) Prey fish escape by sensing the bow wave of a predator. *Journal of Experimental Biology,* 217(24): 4328–4336.

Sticca, E.L., Belbin, G.M., and Gignoux, C.R. (2021) Current developments in detection of identity-by-descent methods and applications. *Frontiers in Genetics,* 12.

Strandburg-Peshkin, A., Twomey, C.R., Bode, N.W., Kao, A.B., Katz, Y., Ioannou, C.C., Rosenthal, S.B., Torney, C.J., Wu, H.S., Levin, S.A., and Couzin, I.D. (2013) Visual sensory networks and effective information transfer in animal groups. *Current Biology*, 23(17): 709-711.

Streelman, J.T., and Albertson, R.C. (2006) Evolution of novelty in the cichlid dentition. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution,* 306(3): 216–226.

Streelman, T.J., and Danley, P.D. (2003) The stages of vertebrate evolutionary radiation. *Trends in Ecology & Evolution,* 18(3): 126–131.

Stroud, J.T., and Losos, J.B. (2016) Ecological opportunity and adaptive radiation. *Annual Review of Ecology, Evolution, and Systematics,* 47(1): 507–532.

Struck, T.H., Feder, J.L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V.I., Kistenich, S. et al. (2018) Finding evolutionary processes hidden in cryptic species. *Trends in Ecology & Evolution,* 33(3): 153–163.

Sugawara, T., Terai, Y., Imai, H., Turner, G.F., Koblmüller, S., Sturmbauer, C., Shichida, Y., Okada, N., 2005. Parallelism of amino acid changes at the RH1 affecting spectral sensitivity among deep-water cichlids from Lakes Tanganyika and Malawi. *Proceedings of the National Academy of Sciences of the USA*, 102(15): 5448–5453.

Suli, A., Watson, G.M., Rubel, E.W., and Raible, D.W. (2012) Rheotaxis in larval zebrafish is mediated by lateral line mechanosensory hair cells. *PLOS One,* 7(2): e29727.

Sulloway, F.J., and Kleindorfer, S. (2013) Adaptive divergence in Darwin's small ground finch (Geospiza fuliginosa): divergent selection along a cline. *Biological Journal of the Linnean Society,* 110(1): 45–59.

Svardal, H., Quah, F.X., Malinsky, M., Ngatunga, B.P., Miska, E.A., Salzburger, W., Genner, M.J, Turner, G.F., and Durbin R. (2020) Ancestral hybridization facilitated species diversification in the Lake Malawi cichlid fish adaptive radiation. *Molecular Biology and Evolution* 37(4): 1100–1113.

Takahashi, T. (2002) Systematics of the tribe Trematocarini (Perciformes: Cichlidae) from Lake Tanganyika, Africa. *Ichthyological Research,* 49(3): 253–259.

Takahashi, T., and Koblmüller, S. (2011) The adaptive radiation of cichlid fish in Lake Tanganyika: a morphological perspective. *International Journal of Evolutionary Biology,* 2011(5): e620754.

Takayama, K., Crawford, D.J., López-Sepúlveda, P., Greimler, J., and Stuessy, T.F. (2018) Factors driving adaptive radiation in plants of oceanic islands: a case study from the Juan Fernández Archipelago. *Journal of Plant Research,* 131(3): 469–485.

Takeda, M., Kusumi, J., Mizoiri, S., Aibara, M., Mzighani, S.I., Sato, T., Terai, Y., Okada, N., and Tachida, H. (2013) Genetic structure of pelagic and littoral cichlid fishes from Lake Victoria. *PLOS One,* 8(9): e74088.

Takuno, S., Miyagi, R., Onami, J., Takahashi-Kariyazono, S., Sato, A., Tichy, H., Nikaido, M., Aibara, M., Mizoiri, S., Mrosso, H.D.J., Mzighani, S.I., Okada, N., and Terai, Y. (2019) Patterns of genomic differentiation between two Lake Victoria cichlid species, *Haplochromis pyrrhocephalus* and *H.* sp. 'macula'. *BMC Evolutionary Biology,* 19(1): 68–74.

Tang, W., Davidson, J.D., Zhang, G., Conen, K.E., Fang, J., Serluca, F., Li, J., Xiong, X., Coble, M., Tsai, T., Molind, G., and Fishman, M.C. (2020) Genetic control of collective behavior in zebrafish. *iScience*, 23(3): 100942.

Tarby, M.L., and Webb, J.F. (2003) Development of the supraorbital and mandibular lateral line canals in the cichlid, *Archocentrus nigrofasciatus*. *Journal of Morphology,* 255(1): 44–57.

Terai, Y., Miyagi, R., Aibara, M., Mizoiri, S., Imai, H., Okitsu, T., Wada, A. et al. (2017) Visual adaptation in Lake Victoria cichlid fishes: depth-related variation of color and scotopic opsins in species from sand/mud bottoms. *BMC Evolutionary Biology,* 17(8): 200–241.

TerMarsch, H., and Ward, J. (2020) Body-generated hydrodynamic flows influence male–male contests and female mate choice in a freshwater fish. *Animal Behaviour,* 169: 119-128.

Thomas, K., Hansen, T., Brophy, D., Maoiléidigh, N.Ó., and Fjelldal, P.G. (2019) Experimental investigation of the effects of temperature and feeding regime on scale growth in Atlantic salmon *Salmo salar* post-smolts. *Journal of Fish Biology,* 94(6): 896–908.

Thompson, A.B., Allison, E.H., and Ngatunga, B.P. (1996) Distribution and breeding biology of offshore cichlids in Lake Malawi/Niassa. *Environmental Biology of Fishes,* 47(3): 235–254.

Thünken, T., Hesse, S., Bakker, T.C.M., and Baldauf, S.A. (2016) Benefits of kin shoaling in a cichlid fish: familiar and related juveniles show better growth. *Behavioral Ecology*, 27(2): 419–425.

Tien, J.H., Levin, S.A., and Rubenstein, D.I. (2004) Dynamics of fish shoals: identifying key decision rules. *Evolutionary Ecology Research,* 6: 555–565.

Tokita, M., Yano, W., James, H.F., and Abzhanov, A. (2017) Cranial shape evolution in adaptive radiations of birds: comparative morphometrics of Darwin's finches and Hawaiian honeycreepers. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 372(1713): 20150481.

Torati, L.S., Migaud, H., Doherty, M.K., Siwy, J., Mullen, W., Mesquita, P.E.C., and Albalat, A. (2017) comparative proteome and peptidome analysis of the cephalic fluid secreted by *Arapaima gigas* (Teleostei: Osteoglossidae) during and outside parental care. *PLOS One,* 12(10): e0186692.

Travis, J. (1989) The role of optimizing selection in natural populations. *Annual Review of Ecology and Systematics,* 20(1): 279–296.

Trokovic, N., Herczeg, G., McCairns, R.J.S., Ab Ghani, N.I., and Merilä, J. (2011) Intraspecific divergence in the lateral line system in the nine-spined stickleback (*Pungitius pungitius*). *Journal of Evolutionary Biology,* 24(7): 1546–1558.

Turner, G., Ngatunga, B.P., Genner, M.J., 2019. The Natural History of the Satellite Lakes of Lake Malawi. *EcoRxiv*. https://doi.org/10.32942/osf.io/sehdq

Turner, G.F. 1996 *Offshore Cichlids of Lake Malawi*, *First Edition*, Cichlid Press.

Turner, G.F., Robinson, R.L., Ngatunga, B.P., Shaw, P.W., and Carvalho, G.R. (2002) Pelagic cichlid fishes of Lake Malawi/Nyasa: biology, management and conservation. In Cowx, I.G. (Ed.) *Management and Ecology of Lake and Reservoir Fisheries*, John Wiley & Sons. 353–366.

Tweddle, D., and Turner, J.L. (1977) Age, growth and natural mortality rates of some cichlid fishes of Lake Malawi. *Journal of Fish Biology,* 10(4): 385–398.

van Dongen, W.F.D., Wagner, R.H., Moodley, Y., Schaedelin, F.C. (2014) Sex biases in kin shoaling and dispersal in a cichlid fish. *Oecologica,* 176(4): 965–974.

van Netten, S.M. (2006) Hydrodynamic detection by cupulae in a lateral line canal: functional relations between physics and physiology. *Biological Cybernetics,* 94(1): 67–85.

van Netten, S.M., and McHenry, M.J. (2014) The biophysics of the fish lateral line. In Coombs, S., Bleckmann, H., Fay, R.R., and Popper, A.N. (Eds.) *The Lateral Line System*, Springer Handbook of Auditory Research. Springer, New York, USA. 99–119.

van Oosterom, L., Montgomery, J.C., Jeffs, A.G., and Radford, C.A. (2016) Evidence for contact calls in fish: conspecific vocalisations and ambient soundscape influence group cohesion in a nocturnal species *Scientific Reports,* 6: 19098.

van Rijssel, J.C., Moser, F.N., Frei, D., and Seehausen, O. (2018) Prevalence of disruptive selection predicts extent of species differentiation in Lake Victoria cichlids. *Proceedings of the Royal Society B: Biological Sciences,* 285(1871): 20172630.

van Trump, W.J., and McHenry, M.J. (2013) The lateral line system is not necessary for rheotaxis in the Mexican blind cavefish (Astyanax fasciatus). *Integrative and Comparative Biology,* 53(5): 799–809.

van Trump, W.J., Coombs, S., Duncan, K., and McHenry, M.J. (2010) Gentamicin is ototoxic to all hair cells in the fish lateral line system. *Hearing Research,* 261: 42-50.

Vandeputte, M., and Haffray, P. (2014) Parentage assignment with genomic markers: a major advance for understanding and exploiting genetic variation of quantitative traits in farmed aquatic animals. *Frontiers in Genetics,* 5(432).

Vernaz, G., Hudson, A.G., Santos, M.E., Fischer, B., Carruthers, M., Shechonge, A.H., Gabagambi, N.P. et al. (2022) Epigenetic divergence during early stages of speciation in an African crater lake cichlid fish. *Nature Ecology & Evolution*, 10: 1–12.

Via, S. (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367 (1587): 451–460.

Viblanc, V.A., Saraux, C., Tamain, A., Criscuolo, F., Coltman, D.W., Raveh, S., Murie, J.O. and Dobson, F.S. (2022) Measuring fitness and inferring natural selection from long-term field studies: different measures lead to nuanced conclusions. *Behavioral Ecology and Sociobiology,* 76(6): 79–85.

Viertler, A., Salzburger, W., and Ronco, F. (2021) Comparative scale morphology in the adaptive radiation of cichlid fishes (Perciformes: Cichlidae) from Lake Tanganyika. *Biological Journal of the Linnean Society,* 134(3): 541–556.

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

Wagner, M., Bračun, S., Duenser, A., Sturmbauer, C., Gessl, W., and Ahi, E.P. (2022) Expression variations in ectodysplasin-a gene (eda) may contribute to morphological divergence of scales in haplochromine cichlids. *BMC Ecology and Evolution,* 22(1): 28–35.

Wainwright, P.C. (2007) Functional versus morphological diversity in macroevolution. *Annual Review of Ecology, Evolution, and Systematics,* 38(1): 381–401.

Wang, H., Xu, C., Liu, X., Guo, Z., Xu, X., Wang, S., Xie, C., Li, W., Zou, C., and Xu, Y. (2017) Development of a multiple-hybrid population for genome-wide association studies: theoretical consideration and genetic mapping of flowering traits in maize. *Scientific Reports,* 7(1): 40239.

Wang, J.Z., Du, Z., Payattakool, R., Yu, P.S., and Chen, C. (2007) A new method to measure the semantic similarity of GO terms. *Bioinformatics,* 23(10): 1274–1281.

Wark, A.R., and Peichel, C.L. (2010) Lateral line diversity among ecologically divergent threespine stickleback populations. *Journal of Experimental Biology,* 213(1): 108–117.

Wark, A.R., Mills, M.G., Dang, L., Chan, Y.F., Jones, F.C., Brady, S.D., Absher, D.M. et al. (2012) Genetic architecture of variation in the lateral line sensory system of threespine sticklebacks. *G3: Genes|Genomes|Genetics,* 2(9): 1047–1056.

Webb, J.F. (1989) Gross morphology and evolution of the mechanoreceptive lateral-line system in Teleost fishes. *Brain, Behavior and Evolution,* 33(1): 34–53.

Webb, J.F. (1990) Ontogeny and phylogeny of the trunk lateral line system in cichlid fishes. *Journal of Zoology,* 221(3): 405–418.

Webb, J.F. (2014) Morphological diversity, development, and evolution of the mechanosensory lateral line system. In Coombs. S., Bleckmann, H., Fay, R.R., and Popper, A.N. (Eds.) *The Lateral Line System*,. Springer Handbook of Auditory Research. Springer, New York, USA. 17–72.

Webb, J.F. (2014) Lateral line morphology and development and implications for the ontogeny of flow sensing in fishes. In Bleckmann, H., Mogdans, J. and Coombs, S. (Eds.) *Flow sensing in Air and Water*, Springer, Berlin, Heidelberg. 247-270

Webb, J.F., Bird, N.C., Carter, L., and Dickson, J. (2014) Comparative development and evolution of two lateral line phenotypes in Lake Malawi cichlids. *Journal of Morphology* 275(6): 678–692.

Weeg, M.S., and Bass, A.H. (2002) Frequency response properties of lateral line superficial neuromasts in a vocal fish, with evidence for acoustic sensitivity. *Journal of Neurophysiology,* 88(3): 1252–1262.

Wellenreuther, M., Brock, M., Montgomery, J.C., and Clements, K.D. (2010) Comparative morphology of the mechanosensory lateral line system in a clade of New Zealand triplefin fishes. *Brain, Behavior and Evolution,* 75(4): 292–308.

Williamson, D., Jackson, M.J., Banerjee, S.K., Merdaci, O., Thouveny, N., Decobert, M. et al. (1999) Magnetic signatures of hydrological change in a tropical maar-lake (Lake Massoko, Tanzania): Preliminary Results. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy,* 24(9): 799–803.

Wilson, L.A.B., Colombo, M., Hanel, R., Salzburger, W., and Sánchez-Villagra, M.R. (2013) Ecomorphological disparity in an adaptive radiation: opercular bone shape and stable isotopes in Antarctic icefishes. *Ecology and Evolution,* 3(9): 3166–3182.

Windsor, S.P., and McHenry, M.J. (2009) The influence of viscous hydrodynamics on the fish lateral-line system. *Integrative and Comparative Biology,* 49(6): 691–701.

Winkelmann, K., Genner, M.J., Takahashi, T., and Rüber, L. (2014) Competition-driven speciation in cichlid fish. *Nature Communications*, 5(3)

Wisenden, B.D. (2000). Olfactory assessment of predation risk in the aquatic environment. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences,* 355(1401): 1205–1208.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1): 3-36

Yasuaki, T., Tetsuya, H., and Juro, Y. (1989) Scale regeneration of tilapia (Oreochromis niloticus) under various ambient and dietary calcium concentrations. *Comparative Biochemistry and Physiology Part A: Physiology,* 92(4): 605–608.

Yoshida, G.M., and Yáñez, J.M. (2021) Multi-trait GWAS using imputed high-density genotypes from whole-genome sequencing identifies genes associated with body traits in Nile tilapia. *BMC Genomics,* 22(1): 57.

Yoshizawa, M., Goricki, S., Soares, D., and Jeffery, W.R. (2010) Evolution of a behavioral shift mediated by superficial neuromasts helps cavefish find food in darkness. *Current Biology,* 20(18): 1631–1636.

Yoshizawa, M., Jeffery, W.R., van Netten, S.M., and McHenry, M.J. (2014) the sensitivity of lateral line receptors and their role in the behavior of Mexican blind cavefish (*Astyanax mexicanus*). *The Journal of Experimental Biology,* 217(6): 886–895.

Zeyl, J.N., Love, O.P., and Higgs, D.M. (2014) Evaluating gonadosomatic index as an estimator of reproductive condition in the invasive round goby, *Neogobius melanostomus*. *Journal of Great Lakes Research,* 40(1): 164–171.

Zhao, H., Mitra, N., Kanetsky, P.A., Nathanson, K.L., and Rebbeck, T.R. (2018) A practical approach to adjusting for population stratification in genome-wide association studies: principal components and propensity scores (PCAPS). *Statistical Applications in Genetics and Molecular Biology,* 17(6)

Zhao, H., Mitra, N., Kanetsky, P.A., Nathanson, K.L., and Rebbeck, T.R. (2018) A practical approach to adjusting for population stratification in genome-wide association studies: principal components and propensity scores (PCAPS). *Statistical Applications in Genetics and Molecular Biology,* 17(6): 1-12

Zheng, X., Wang, C., Fan, R., and Xie, G. (2018) Artificial lateral line based local sensing between two adjacent robotic fish. *Bioinspiration & Biomimetics,* 13: 016002.

Zheng, X., Wang, M., Zheng, J., Tian, R., Xiong, M., and Xie, G. (2019) Artificial lateral line based longitudinal separation sensing for two swimming robotic fish with leader-follower formation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2019*, 2539-2544.

Zheng, X., Wang, W., Li, L., and Xie, G. (2020) Artificial lateral line based relative state estimation between an upstream oscillating fin and a downstream robotic fish. *Bioinspiration & Biomimetics,* 16(1): 016012.

Zheng, X., Wang, W., Xiong, M., and Xie, G. (2020) Online state estimation of a fin-actuated underwater robot using artificial lateral line system. *IEEE Transactions on robotics*, 36(2): 472-487.

Zhou, X., and Stephens, M. (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics,* 44(7): 821–824.

Zhou, Z., and Liu, Z. (2008) Biomimetic cilia based on MEMS technology. *Journal of Bionic Engineering,* 5: 358-365.