Durham University

# Durham E-Theses

## *Adaptation in the Deep Sea: How Depth Affects Morphology and Protein Function in Coryphaenoides rupestris and its Congeners*

STEEDS, NATASHA,JANE

**How to cite:**

STEEDS, NATASHA,JANE (2020) *Adaptation in the Deep Sea: How Depth Affects Morphology and Protein Function in Coryphaenoides rupestris and its Congeners*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/13841/

**Use policy**

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

# Adaptation in the Deep Sea: How Depth Affects Morphology and Protein Function in *Coryphaenoides rupestris* and its Congeners

Natasha Steeds

Submitted to the Department of Biosciences, Durham University in fulfilment of the requirements for the degree MSCR Biological Sciences C1A009

March 2020

# Material Abstract


Adaptation in the Deep Sea: How Depth Affects Morphology and Protein Function in *Coryphaenoides rupestris* and its Congeners

Natasha Steeds

Although the deep sea is the largest habitat on Earth, there is little understanding of how adaptation and speciation occur across depth gradients. The roundnose grenadier, *Coryphaenoides rupestris*, has demonstrated genetic segregation according to habitat depth across several functional loci. In this study, 139 *C. rupestris* individuals were sampled for morphological analysis, and the protein OBSL1 was modelled to explore the impact of non-synonymous single nucleotide polymorphisms (SNPs) across the species' depth range. It was confirmed that the genetic segregation is mirrored by morphological differences in deeper-living individuals. *C. rupestris* sampled from deeper habitats were smaller and had more slender body forms, along with larger eyes and mouth gapes. Lipid stores in the liver and swimbladder both increased with habitat depth. OBSL1, the modelled muscle protein, was shown to have tighter intra-domain boning in fish from shallower habitats. These changes were linked to changes in foraging strategy and an increased demand for energy conservation in deeper habitats.

Analyses were extended to explore adaptation to even deeper habitats. OBSL1 was modelled for 13 additional *Coryphaenoides* species from abyssal and non-abyssal habitats. These analyses suggested that hydrostatic pressure was the key selection pressure for this protein as habitat depths increase to reach the abyssal zone. This finding was contrasted with those from the single-species analysis of *C. rupestris* to illustrate how selection pressures for adaptation and speciation change across a depth gradient in the deep sea.

# Table of contents

**Material Abstract** ..... **i**

**Table of Contents** ..... **ii**

**Statement of Copyright** ..... **iv**

**Acknowledgements** ..... **v**

**Chapter 1: Introduction and Background** ..... **1**

*Potential for adaptation in the deep sea* ..... 1

*Phenotypic responses to environmental factors* ..... 2

*The study species* Coryphaenoides rupestris ..... 5

*Fishery, conservation and management* ..... 8

*Project rationale and aims* ..... 10

**Chapter 2: Morphometric Analysis of *C. rupestris*** ..... **11**

**Introduction** ..... 11

*The mesopelagic-bathypelagic boundary* ..... 11

*Expected morphological trends in* C. rupestris ..... 12

**Methods** ..... 18

*Measurement and dissection protocol* ..... 19

*Otolith preparation and sectioning* ..... 21

*Statistical analysis* ..... 23

**Results** ..... 27

*Demographic and general trends* ..... 27

*Single trait tests* ..... 30

*Principal component analysis* ..... 36

*Clustering analyses* ..... 41

**Discussion** ..... 43

*Demographic and general trends* ..... 43

*Body size* ..... 44

*Body shape* ..... 46

*Energy storage* ..... 49

*Conclusion* ..... 51

**Chapter 3: Changes to Protein Function Across the Meso-Bathypelagic Boundary in *C. rupestris*** ..... **53**

**Introduction** ..... 53

*Known protein adaptations to life in the deep sea* ..... 53

*Initial exploration of potential study proteins* ..... 56

*OBSL1 and its function* ..... 59

**Methods** ..... 62
*Initial analysis* ..... 62
*OBSL1 size analysis* ..... 63
*SNP modelling* ..... 64
**Results** ..... 66
*Initial analysis* ..... 66
*OBSL1 size analysis* ..... 68
*SNP modelling* ..... 69
**Discussion** ..... 74
**Chapter 4: Beyond *C. rupestris*: Changes to Protein Structure and Function Across the Abyssal Boundary in *Coryphaenoides*** ..... **82**
**Introduction** ..... 82
*The abyssal zone* ..... 82
*Study protein and rationale* ..... 83
**Methods** ..... 85
**Results** ..... 87
**Discussion** ..... 93
**Chapter 5: Synthesis and Conclusion** ..... **99**
*Relevance of ecotype segregation to conservation* ..... 99
*Linking morphology with protein function: the role of locomotion* ..... 100
*Further considerations* ..... 103
*Conclusion* ..... 104
**Appendices** ..... **106**
**Bibliography** ..... **121**
**Webography** ..... **128**

## Statement of Copyright

*The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.*

# Acknowledgements

I am grateful to Dr Thomas Regnier for his help in sectioning the otoliths as discussed in Chapter 2, as well as his contribution to the analysis of morphological data for that chapter. I also thank Dr Georgios Gkafas for his assistance in navigating genomic datasets, and his work creating multiple data sheets which allowed me to carry out my analyses.

My thanks also go to my supervisors, Prof. Rus Hoelzel and Dr Ehmke Pohl, whose guidance and expertise were invaluable during the course of this project.

# Chapter 1: Introduction and Background

*Potential for adaptation in the deep sea*

The deep sea is the largest and least understood habitat on Earth. Widely defined as the environment deeper than 200 m (Gage and Tyler, 1992; Priede, 2017), the deep sea comprises 63% of global surface area, ~95% of ocean volume, and >90% of the Earth's habitable volume (Childress, 1995; Smith et al., 2008; Thurber et al., 2014; Baco et al., 2016). This is an environment characterised by harsh conditions: average temperatures are below 4ºC, and average hydrostatic pressure is 400 atm (Danovaro et al., 2014). Historically, because of these extreme environmental conditions and the widely-held assumption that local variation in surface-level productivity would not affect deeper living communities, the deep sea was long presumed to be a homogenous habitat with poor biodiversity (Wilson and Hessler, 1987; Rex et al., 1993). However, the discovery of staggeringly high biodiversity in the deep sea (Hessler and Sanders, 1967; Ramirez-Llodra et al., 2010; Thurber et al., 2014), as well as latitudinal and regional variations in species turnover and biodiversity (Rex et al., 1993; Culver and Buzas, 2000; Rex, et al., 2000; Levin et al., 2001; Glover et al., 2002; Lambshead et al., 2002), has led to the acceptance of the deep sea as a diverse and heterogeneous environment, about which we still know very little.

While topographical and geological features such as hydrothermal vents and mid-ocean ridges are highly influential in determining adaptation, speciation and biodiversity patterns (Van Dover et al., 2002), perhaps the most striking environmental drivers of such processes in the deep sea are associated with depth. Pressure, salinity, oxygen concentration, dissolved nutrient availability, food abundance, light intensity, and temperature all change to varying degrees as a function of depth in the ocean, and any one of these factors may drive changes in adaptation, abundance, and diversity of life. Hydrostatic pressure increases by ~100 kPa with every 10 m of depth (Childress, 1995). Heat from the Sun is absorbed by only the first

few metres of ocean surface, and temperatures gradually decline with depth to reach ~1˚C in the abyssal zone, ≥ 3,000 m (Gage and Tyler, 1992). Light and oxygen levels also decrease with depth. Light becomes insufficient for photosynthesis deeper than 100-150 m, and reaches zero at around 1,000 m (Ryther, 1956; Priede, 2017). Oxygen levels decrease below the photic zone due to the sinking and degradation of organic materials, reaching a minimum layer at 1000 m or so (Jewell, 1994). Below this minimum layer, oxygen concentration increases again as a result of decreasing biological oxygen demand and other environmental factors (Wyrtki, 1962; Rogers, 2000).

Salinity also changes with depth, peaking just below the surface before reducing to a minimum at around 700 m, increasing slightly to stabilise to an almost isohaline concentration deeper than 1000 m (Emery and Dewar, 1982). Seasonal and regional salinity fluctuations are mainly restricted to the ocean surface layer (Foltz and McPhaden, 2008). Food availability also changes with depth: though decreasing animal abundance and lack of primary productivity reduces quantities of large food items in the deep sea (Childress, 1995; Collins et al., 2005), dissolved nutrients from shallower waters provide some energy, while sinking carcasses provide occasional scavenging opportunities (Priede, 2017). It is clear that the deep sea offers several crucial environmental gradients which could facilitate adaptive differentiation between populations at different depths. However, because there are so many environmental factors which change with depth, it is difficult to determine which might be the most crucial selection pressures for deep sea adaptation. The question remains as to which morphological and physiological characters are likely to undergo adaptive change in response to these selection pressures.

*Phenotypic responses to environmental factors*

There are a number of phenotypic trends that have been observed in deep sea species which can be used to intuit the relative importance of different deep sea environmental gradients for adaptation and speciation. Some of these trends involve the coordination of multiple

morphological characters; for example, Neat and Campbell (2013) noted that deeper living fish species tended to have more elongate, slender body plans compared to shallower species. While the selection pressure for this adaptation was not clearly identified, the authors suggested links to altered foraging strategies in deeper species as a result of decreasing food availability and demands on energy use. Highly specific physiological trends have also been discovered, including changes in osmolyte concentrations in the tissues of deep sea species. Trimethylamine oxide (TMAO) has consistently been shown to increase in body tissues of teleosts, skates and crustaceans with increasing habitat depth, while urea concentration decreases with depth (Gillett et al., 1997; Yancey et al., 2002). It has been suggested that these changes allow for improved stability in enzymes which would become less functional at the higher hydrostatic pressures experienced as depth increases (Yancey et al., 2004). Other phenotypic trends, such as increasing lipid content in cell membranes as a response to increasing hydrostatic pressure (DeLong and Yayanos, 1985), have been noted, however none are more thoroughly documented than the changes to metabolism at increasing depths.

Metabolic rate has been shown to decrease with depth across taxa (summarised in Childress, 1995), and there is considerable evidence for metabolic enzymes undergoing intrinsic changes in deep sea animals. Although the trend is widely accepted, there is some contention over the key selection pressures which drive this adaptation. One explanation championed by Childress (1995) is that as light levels decrease with depth, less importance is placed on locomotory performance for foraging and vertical migrations. Indeed, several studies in *Sebastes spp.* have demonstrated links between rhodopsin evolution and species depth, giving strong support to the argument of light availability as a driver in deep sea adaptation (Sivasundar and Palumbi, 2010; Shum et al., 2014). Alternatively, Vetter and Lynn (1997) suggest that declining food availability is responsible for the metabolic trend, citing that artificially increasing food availability in metabolically depressed deep sea species can

reverse the effect, increasing metabolic rate (Sullivan and Somero, 1983; Yang and Somero, 1993; Vetter et al., 1994).

Some of the most striking data concerning metabolic depression in deep sea species and its potential mechanisms concern the metabolic enzymes themselves. Seibenaller and Somero (1982) show that muscular kinase and dehydrogenase activity was significantly lower in the deeper living of two *Sebastolobus* species with similar morphologies and ecology. Similarly, low activity in muscular glycolytic enzymes including lactate dehydrogenase (LDH) was found in five species of deep sea macrourid fishes compared to shallow living species (Seibenaller et al., 1982), and a comparable decrease in muscle LDH activity was observed in the dover sole (*Microstomus pacificus*; Vetter et al., 1994). Interestingly, Vetter and Lynn (1997) showed that while activity of LDH and several other metabolic enzymes decreased as average habitat depth of *Sebastolobus spp.* increased, higher enzyme activitiy levels were maintained when the generally shallow living *S. alascanus* migrated to greater depths. This suggests that lowered metabolic activity at greater depths is an adaptive response to continuous living in the deep sea, as opposed to a flexible functional response to changing environmental conditions. Local changes in enzymes could have significant effects on this metabolic trend, and these differences themselves might be adaptations to deep sea living. Metabolic activity differences have been attributed to increased activation free energy and enthalpy characteristics in the enzymes of deep living fish species, making them less functional (Somero and Seibenaller, 1979). Somero (1990) linked the trend of lowered enzyme activity in deeper waters to reduced pressure sensitivity in the enzymes in question, allowing them to function, albeit less efficiently, at increased hydrostatic pressures.

It is clear from the literature that some of the most significant adaptations to life in the deep sea concern reduced metabolism at greater depths in relation to overall energy use, metabolic enzyme activity, and intrinsic enzyme function. The most significant environmental pressures which impact the observed trends appear to be the increase in hydrostatic pressure

associated with increasing depth and the constraints placed on energy use by diminishing resource and light abundance. In this thesis, I will explore the dynamics and mechanisms of these selection pressures and resultant adaptations within one genus of teleosts, examining one species in particular: the roundnose grenadier, *Coryphaenoides rupestris*. It is my aim for this work to shed light on how adaptation and speciation might occur with increasing depth in the oceans, and to use this information to articulate goals for appropriate management of deep sea fishes.

*The study species Coryphaenoides rupestris*

The deep sea macrourid *Coryphaenoides rupestris* is the focus of this project. The genus *Coryphaenoides* is a diverse group known as grenadiers or rattails, composed of 66 species living across a vast depth gradient (Cohen et al., 1990; Gaither et al., 2016a). They have slender, elongate body plans with long tails which comprise the majority of their body length. They can reach over 120 cm in length, but are generally below 60 cm (Cohen et al., 1990). While the shallowest species can be found within the euphotic zone at ~110 m (Gaither et al., 2016a), other species live as deep as the hadal zone, up to ~7000 m (*C. yaquinae*; Linley et al., 2016). The total depth range for *Coryphaenoides* is huge, but most species live between 700 and 2000 m (Cohen et al., 1990). Geographically, they are found across a large range from polar to tropical seas, and while most species have set ocean ranges, two species are considered circumglobal (*C. armatus* and *C. rudis*; Gaither et al., 2016a; Gaither et al., 2016b). Recent studies in this group have led to improved understanding of how the abyssal zone might be colonised by shallower living species (Gaither et al., 2016a), and the diverse depth and geographical ranges of *Coryphaenoides spp.* make it an ideal genus to explore depth-associated adaptation and speciation.

*C. rupestris*, the focal species of this project, is in many ways typical of the genus, with an elongate body and long, tapering tail. It has a rounded, blunt nose which extends beyond the snout, a compressed posterior body shape, and large egg-shaped eyes (Figure 1). Sharp,

villiform teeth are present in both the upper and lower jaw (Atkinson, 1995) This is a benthopelagic species, living close to the sea bottom, feeding on a mixture of cephalopods, copepods and other teleost fishes. Interestingly, many of these prey species are pelagic (Bergstad et al., 2010), and there is some evidence that *C. rupestris* might participate in diel vertical migration, rising in the water column to access food resources (summarised in Atkinson, 1995). There is no consensus on these data however, and it remains equally possible that *C. rupestris* takes advantage of the downwards vertical migrations of its pelagic prey, rising only slightly to meet them, as opposed to itself migrating to pelagic areas to hunt (Mauchline and Gordon, 1991; Lorance and Trenkel, 2006). Relative quantities of prey items in the diet have been shown to change as fish mature, with younger individuals relying more heavily on cephalopods, and the percentage of fish and copepods in the diet increasing with maturity and size (Bergstad et al., 2010). Their low metabolism and large energy stores, plus their body plan designed for slow cruising, have been used to suggest that these fish engage mostly in scavenging or 'sit and wait' predation as opposed to active hunting (White et al., 2010). This is supported by behavioural data from Lorance and Trenkel (2006), who observed *C. rupestris* individuals mostly drifting, swimming or station holding, as opposed to engaging in active swimming or pursuit.

Figure 1: *C. rupestris* individual. Photograph taken during dissection for this project (Chapter 2)

As is common in deep sea fish, growth is very slow for this species, with males and females only reaching sexual maturity at 8.5 and 11 years respectively (Bergstad, 1990). Slow life

histories are translated to long lifespans and large sizes at maturity, with individuals recorded as old as 72 years, and at lengths of up to 130 cm (Atkinson, 1995; Bergstad, 1990). This is a batch spawning species, with females capable of producing up to 70,000 oocytes at a time (Allain, 2001). Relatively little is known about the process of *C. rupestris* reproduction; while there is evidence for a peak Autumn breeding season, and behavioural data suggesting that individuals aggregate into dense clusters for spawning (Bergstad, 1990; Neat, 2017), oocyte development and release does not appear to be synchronised across individuals (Allain, 2001). *C. rupestris* is likely the most abundant and well-studied of the macrourids (Atkinson, 1995), and can be found across continental slopes in the Northeast and Northwest Atlantic and mid-Atlantic ridge (White et al., 2010) at depths between 180 and 2600 m (Cohen, 1990).

Recent genetic studies have brought increased focus to *C. rupestris* as a study species for exploring adaptation in the deep sea. White et al. (2010) examined 16 hypothetically neutral microsatellite loci in 417 individuals from the North Atlantic, and found that along with a significant decrease in genetic diversity, there was evidence for local selection linked to a microsatellite DNA locus labelled Crup7 in fish from habitats deeper than 1200 m. Though a gene under selection linked to Crup7 was not identified, this study determined the potential for genetic structuring according to depth in this species. Gaither et al. (2018) built upon this work and generated an annotated reference genome for *C. rupestris*, against which they compared 60 additional *C. rupestris* genomes from a range of habitat depths between 750 and 1800 m. This comparison allowed them to demonstrate differentiation between the genomes which originated from different depths. They were able to show disruptive selection according to habitat depth in the form of non-synonymous single nucleotide polymorphisms (SNPs) in 6 functional loci, however found no evidence of such differentiation at neutral loci. They used this evidence to conclude that assortative mating at different depths was not occurring or not yet detectable in this species, but rather that mating

was effectively random, and that individuals with certain genotypes would migrate to greater depths as they matured, thereby creating a distinct differentiation in ecotype according to depth. These two studies provided clear evidence that there was some level of genetic structuring associated with depth in this species, however it remained unclear what the phenotypic consequences of this structuring were, and indeed what selection pressures were driving the changes. It is my intention to build upon these studies and elucidate what phenotypic changes are present in this species with changing depth, and how they might have arisen.

*Fishery, conservation and management*

In addition to exemplifying ecotype segregation across its large depth range, *C. rupestris* is also an appropriate species to study because of its relevance to commercial fisheries. Commercial exploitation of *C. rupestris* began in the 1960s and this species has remained a key target species for deep water fisheries in both the Northwest and Northeast Atlantic, its fatty liver and pleasant tasting meat popularising it as a food source (Atkinson, 1995; Iwamoto, 2015). Specific quotas for *C. rupestris* catch were only put in place in 1972 in the Northwest Atlantic, and were very high, despite the fact that very little was known about the demographics and ecology of the species at the time (Atkinson, 1995; Haedrich et al., 2001). Catches at this time were enormous, with a peak in the 1970s of approximately 80,000 tonnes annually (Koslow et al., 2000), but since dropped off dramatically, to the point where the Northwestern fishery essentially shut down in the 1990s (Haedrich et al., 2001; Iwamoto, 2015). A fishery continues in the Northeast Atlantic, though catches are very small. Most of the fishing pressure placed on this species today appears to be from bycatch.

The impact of the overexploitation of this fish stock in the 1970s is still being felt. At present, *C. rupestris* is endangered in the Northwest Atlantic according to the Committee on the Status of Endangered Wildlife in Canada (COSEWIC; Baker et al., 2009; Simpson et al., 2011). When Devine et al. (2006) applied criteria from the International Union for

Conservation of Nature (IUCN) to *C. rupestris*, the species qualified as endangered, meaning over 70% population decline was observed over 10 years as a result of fishing practices. Indeed, according to the IUCN's current red list, *C. rupestris* is a critically endangered species (Iwamoto, 2015). The slow growth and K-selected life history of *C. rupestris* mean that recovery from such significant population reduction is not certain (Koslow et al., 2000; Baker et al., 2009). Recruitment in *C. rupestris* is not consistent, meaning recovery is impossible to properly predict, and may be more sensitive to stochastic population events. Bergstad et al. (2014) demonstrated that only one large recruitment event had occurred in Skagerrak *C. rupestris* in the last three decades. Though this population is more isolated than most, and may therefore face slightly different demographic dynamics, the potential for highly sporadic recruitment events is likely to be consistent across the species, meaning that the risk *C. rupestris* faces from fishing exploitation is severe.

Several management practices have already been suggested and implemented for this species. The International Council for the Exploration of the Sea (ICES) advised that *C. rupestris* catch be reduced to zero to avoid further unsustainable exploitation of this already sensitive fish stock (ICES, 2016). The IUCN proposed improved monitoring of population dynamics and catch reduction, as well as implementation of larger net mesh sizes to reduce bycatch of juveniles (Iwamoto, 2015). Though these strategies are sound and should be applied, they do not consider how management should be approached across the whole depth range that *C. rupestris* inhabits. Historically, fishing effort has been concentrated at the top half of this species' depth range, generally remaining shallower than 1500 m (Atkinson, 1995; Neat and Burns, 2010). It has not been established how selectively removing shallower living individuals might have affected the overall population. In the light of the genetic information now available, which demonstrates the potential for ecotype segregation according to depth (Gaither et al., 2018), it is important now to explore the dynamics of this segregation, and to use this information not only to show how fishing practices might impact

population dynamics in unexpected ways, but also to inform future management of this stock.

*Project rationale and aims*

I believe that this project will bring an improved understanding of how environments stratified by depth in the deep sea might bring about adaptation, and indeed speciation. This is a topic about which we still know surprisingly little, and so improved understanding of how environments in the deep sea facilitate adaptation and evolutionary change is crucial. Furthermore, the use of the critically endangered *C. rupestris* as a study species will hopefully allow me to demonstrate how fishing practices might impact deep sea populations, and inform future management of this species.

I plan to build on the genetic studies carried out by White et al. (2010) and Gaither et al. (2018), which suggested genetic segregation according to depth in *C. rupestris*, but were unable to identify phenotypic consequences to these changes, or selection pressures that might have influenced the differentiation. Using a combination of morphometric, bioinformatic and proteomic approaches, I will attempt to link phenotypic changes which occur with depth in this species to the genetic changes that have already been observed. This will allow me to propose key environmental factors in the deep sea which might act as selection pressures for this species, and perhaps to identify the impact of historical fishing on this population. Not only will understanding this segregation improve understanding of adaptation according to depth, but it will allow me to suggest how proper management should be implemented to reduce further damage to the population by fishery. In addition to this, I will investigate proteomic changes between different members of the *Coryphaenoides* genus in an attempt to demonstrate how adaptation and speciation might occur across the abyssal boundary. This will further improve our understanding of how the selection pressures at work in the deep sea impact adaptation and speciation according to depth.

# Chapter 2: Morphometric Analysis of *C. rupestris*

## Introduction

*The mesopelagic-bathypelagic boundary*

Gaither et al. (2018) recently suggested that genetic segregation at functional loci in *C. rupestris* is associated with the boundary between the mesopelagic and bathypelagic layers in the deep ocean. Classically, the mesopelagic has been categorised as the region between the epi- and bathypelagic, at 200-1000 m. This is a definition based on light availability: the upper boundary with the epipelagic is where light becomes insufficient for photosynthesis, and the lower boundary with the bathypelagic is the maximum depth of solar light penetration in clear water (Priede, 2017). The penetration of low-level light which characterises the mesopelagic gave rise to its colloquial name, "the twilight zone". The bathypelagic, between 1000 and 3000 m, is the largest layer of the deep pelagic and the largest ecosystem on the planet, comprising almost 75% of ocean volume (Ramirez-Llodra et al., 2010). This zone is characterised by a complete lack of solar light.

In order to understand how and why genetic segregation across the meso-bathypelagic boundary might have occurred in *C. rupestris*, it is important to establish the key changes which take place across this gradient. Temperature decreases with depth in a permanent thermocline through the mesopelagic, but is fairly consistent in the bathypelagic (Priede, 2017). The oxygen minimum layer is at a similar depth to the meso-bathypelagic boundary at 1000 m. Light availability decreasing to zero at the meso-bathypelagic boundary influences biomass limitations at different depths. Without the potential for photosynthesis, many deep sea fauna must rely on food sinking or advection from the epipelagic for food (Ramirez-Llodra et al., 2010). Food particles are less likely to penetrate to greater depths: indeed, Buesseler et al. (2007) revealed that only 0.5-2% of net primary production from the epipelagic will reach the sea floor deeper than 2000 m. Food therefore becomes more

scarce in the bathypelagic compared to the mesopelagic, resulting in lower faunal biomass and productivity.

Faunal biodiversity increases with depth, reaching a maximum at 2000-3000 m (Ramirez-Llodra et al., 2010). Though perhaps surprising considering the diminished biomass at depth, greater biodiversity can be linked to the long-term stability of the bathypelagic, often considered the oldest ecosystem in the world (Sutton et al., 2010). Characteristics of deep sea fauna also change across the meso-bathypelagic gradient. Species in the mesopelagic often perform diel vertical migrations at night to the epipelagic to feed, while it has been suggested that bathypelagic organisms show greater fidelity to habitat depth (Reygondeau et al., 2018). Drazen and Sutton (2017) concluded that bathypelagic species were more likely to be generalist feeders based on their large mouth gapes and reduced musculature. They also suggested that mesopelagic species could rely more heavily on visual cues for predation due to the presence of low level light in this zone.

It is important to note that the boundary between the mesopelagic and bathypelagic is not set finitely at 1000 m in all areas of the ocean. All factors which can be used to describe these two pelagic zones vary both seasonally and regionally. Reygondeau et al. (2018) recently showed, using computer modelling of environmental climatology data, that the vertical coverage of the mesopelagic zone may actually vary between 50 and 2300 m according to global biogeography. This being said, as *C. rupestris* lives between 180 and 2600 m, any definition of the meso-bathypelagic boundary would be applicable to the observed genetic segregation. Here, the boundary of 1500 m is investigated in particular detail, as this was the depth identified by Gaither et al. (2018) as showing genetic segregation according to the meso-bathypelagic boundary.

*Expected morphological trends in* C. rupestris

Here, I build on the previous studies by White et al. (2010) and Gaither et al. (2018), which showed genetic segregation in *C. rupestris* across a depth gradient. I focus on how

morphology of *C. rupestris* individuals might be expected to change alongside this genetic structuring across the meso-bathypelagic gradient. There are several noted morphological trends which take place across this depth gradient in the ocean, however most of these come from multi-species studies; there are few studies which examine morphological changes with depth in a single species. Nevertheless, the patterns which have been observed across species are worthy of note in this context because they show how adaptation to certain depths might impact morphological traits, and these principles may extend to within-species variation.

Body size has a well-explored, but complicated relationship with depth in the ocean. Both dwarfism and gigantism have been observed in the deep sea benthos, and these conflicting patterns have been difficult to assimilate into a single theory of body size adaptation to depth (Ramirez-Llodra et al., 2010). Both trends, however, have been linked to food availability decreasing with depth. Large body sizes have a lesser energy demand per unit mass, and so are metabolically more efficient than smaller body sizes (Peters, 1986). Large body sizes also confer competitive advantages, so could be favoured as food becomes limited with depth. This was the suggested mechanism for the increasing size of eight benthic gastropod species with habitat depth (Rex and Etter, 1998), and similar trends have been observed in multiple crustacean groups as well (Wilson et al., 2015). More commonly observed by far, however, is the trend for decreasing body size with depth. This is predicted by Thiel's size-structure hypothesis (1975; 1979), which suggests that despite higher relative metabolic cost, the lower individual food demands of smaller individuals would promote a decrease in body size as food availability becomes limited with depth. Reducing individual food intake with small body sizes permits a larger number of conspecifics to be supported, allowing a sufficient total population size for sustaining the population over subsequent generations (Kaariainen and Bett, 2006; Ramirez-Llodra et al., 2010). This trend has been documented across a wide range of deep sea taxa. Kaariainen and Bett (2006) showed that average body size of large and small macrofauna, mesofauna and meiofauna all decreased with depth when

comparing samples from 150 and 1600 m. Rex et al. (2006) also supported this trend, reporting body size decreases across meiofauna, macrofauna and megafauna with increasing depth.

Looking at deep sea fish species in particular, conflicting trends can also be observed in body size. Collins et al. (2005) showed with trawl data of 76 demersal fish species from between 800 and 4800 m that the mean body size of scavenging fish species increased with depth, while size of non-scavenging fish decreased. They suggested that larger sizes in scavengers permitted higher swimming speeds and endurance, as well as improved mass-specific metabolic efficiency, allowing improved survival on sporadic scavenged food items. The partitioning of body size trends observed here according to feeding guild supports ideas such as Thiel's size-structure hypothesis which link body size to feeding ecology and food limitation with increasing depth. It is clear that a species' overall ecology and niche are important determinants of what biological patterns they fit in to. Taking the trends observed by Collins et al. (2005), we can expect *C. rupestris*, as a mainly non-scavenging species, to fall into the pattern of decreasing body size with depth. Whether this inter-species trend will extend to within-species variation, however, remains to be seen.

Other whole-body morphological trends have been observed in deep sea fishes. Neat and Campbell (2013) observed a trend for increased body elongation at greater habitat depths in *c.* 266 fish species found across the meso- and bathypelagic in the Northeast Atlantic. This change was linked to improved efficiency of anguilliform swimming. This form of locomotion is observed in long, slender fish such as eels, and involves wave-like side-to-side undulations across the body's length. Anguilliform swimming is more efficient in more viscous water, and water viscosity increases slightly with hydrostatic pressure and therefore depth (Likhachev, 2003). Furthermore, improved swimming efficiency no doubt becomes crucial when the environment is food limited, as occurs with increasing depth in the ocean (Drazen and Sutton, 2017). Again, if inter-species trends in depth adaptation can be mirrored

in intra-species variation, we might expect deeper-living *C. rupestris* individuals to display a narrower, more elongate body shape than those from shallower environments.

More specific morphological trends have also been observed across depth gradients in the deep sea, and these too could be reflected in variation in *C. rupestris* morphology. Another pattern associated with feeding ecology is the trend for increased mouth gape size with depth, allowing bathypelagic species whose food is more limited to fill a more generalised predatory niche and access a wider range of food resources (Drazen and Sutton, 2017). For example, in Ebeling and Caillet's study of 31 meso- and bathypelagic lanternfish and bigscale species (1974), it was shown that mouth size increased among bathypelagic species independent of body size in both taxa. Associated with increasing mouth size with habitat depth is the trend for reduction in gill rakers on the first gill arch. The size of the gap between gill rakers has been shown to vary functionally: species with closer spaced gill rakers are more likely to feed on small food items such as plankton (Magnuson and Heitz, 1971). Ebeling and Caillet (1974) showed that mesopelagic lanternfish and bigscale species had significantly more gill rakers, and therefore smaller spaces between rakers, than bathypelagic species, supporting the suggestion that bathypelagic species tend to be adapted to broad, generalist diets in response to low food availability, where mesopelagic species are more able to access more specialised or smaller food sources. This trend for generalism in the bathypelagic as expressed by increasing mouth size and decreasing gill raker number might be reflected in *C. rupestris*.

Eye size might also be expected to change across the habitat range of *C. rupestris*, as this species occupies both the dimly lit mesopelagic, and the completely dark bathypelagic. Eyes require a great deal of energy in both development and neural maintenance (Niven, 2015), and the eyes of *C. rupestris* are large and so costly to maintain (Atkinson, 1995). In the mesopelagic, the cost of having large eyes might be outweighed by the benefits of improved hunting ability and predator detection, however the same may not be true in the dark

bathypelagic. Studies of eye size have shown that eyes tend to be smaller in bathypelagic species. The trend of decreasing eye size with increasing depth could therefore be expected in this species. Smaller eye size would be more energy efficient in the bathypelagic, where food is scarce.

The physiology of the swimbladder plays an important role in adaptation to depth in deep sea teleosts, and so should be explored in *C. rupestris*. The swimbladder is a gas-filled organ essential for buoyancy and therefore energy-efficient swimming. In *C. rupestris*, as in many deep sea teleosts, the swimbladder contains lipids comprised mostly of cholesterol and phospholipids, which are saturated with oxygen bubbles providing buoyancy (Phleger, 1991; 1998). These oxygen bubbles are delivered to the swimbladder by the Root effect, which is a phenomenon observed in fish haemoglobin wherein decreasing pH and increased blood carbon dioxide concentration decrease the haemoglobin's oxygen carrying capacity, resulting in oxygen dissociation and offloading into the swimbladder (Pelster and Weber, 1991). There has been some suggestion that the Root effect becomes limited at high hydrostatic pressures and is therefore not responsible for swimbladder inflation in deep sea species. This led to the belief that the lipids in the swimbladder themselves might aid oxygen dissociation and swimbladder inflation (Scholander, 1954; Phleger, 1998). If this were the case, a greater mass of lipids could be expected with increasing habitat depth in *C. rupestris*. Scholander (1954), demonstrated that *C. rupestris* haemoglobin oxygen saturation is reduced in the presence of lactic acid at atmospheric pressures of over 150 atm, suggesting the Root effect is still functional at high pressure, however this pressure is experienced only deeper than around 1500 m in the deep sea. The Root effect may become less functional below this level, where hydrostatic pressure increases, necessitating increased lipid content in the swimbladder to facilitate oxygen dissolution.

Swimbladder lipid mass may be expected to increase with habitat depth in *C. rupestris* due to other factors as well. The partial pressure gradient between the swimbladder gases and

environment increases with increasing habitat depth and hydrostatic pressure, resulting in the potential for diffusional loss of swimbladder gases and the buoyancy they confer (Pelster, 1997). Swimbladder lipids may serve as a barrier to back-diffusion of oxygen, meaning that increased lipid content in the swimbladder with increasing habitat depth could improve buoyancy and minimise energy loss to swimming (Phleger, 1998). The buoyancy conferred by the gases in the swimbladder also decreases with increasing habitat depth because of the way that the relative density of swimbladder gases increase with increasing external hydrostatic pressure (Pelster, 1997). Lipids have lower densities than water, so it would stand to reason that as hydrostatic pressure increases with habitat depth and the buoyancy conferred by swimbladder gases is reduced, swimbladder lipid content also increases. A combination of these factors allows us to expect a greater mass of lipid in the swimbladders of deeper living *C. rupestris* individuals.

Finally, changes in liver mass might be expected with increasing depth in *C. rupestris*. Energy storage in three other *Coryphaenoides* species (*C. armatus, C. yaquinae* and *C. acrolepis*) was shown to be mainly in the form of neutral lipids in the liver (Drazen, 2002), so we can expect that the same is true in *C. rupestris*, and that the mass of the liver reflects lipid energy stores. Large lipid energy stores in deep sea teleosts are associated with long periods of fasting (Musick and Cotton, 2015); indeed, Smith (1978) showed that the liver lipid stores of the highly food limited abyssal grenadier *C. armatus* were sufficient for survival for *c.* 186 days without feeding. It stands to reason therefore that *C. rupestris* individuals from the bathypelagic, who must survive longer periods between meals, would allocate more energy to storage in the liver, resulting in greater liver masses relative to body size, than their mesopelagic conspecifics.

Here I have outlined some of the documented morphological tends associated with increasing habitat depth, and how they can be expected to be matched in *C. rupestris* across its depth range. A concise summary of the expected patterns can be found in Table 1. It was

important to thoroughly test the relevance of these trends in *C. rupestris*, and whether any additional adaptive or neutral changes in morphology were occurring, as these could further reflect the ecotype segregation that this species is experiencing according to depth. A large set of external and internal measurements of body features and organs was therefore obtained through measurement and dissection, in the hopes of revealing any additional morphological trends across this species' depth range.

Table 1: Summary of morphological features and hypotheses for how they will change with increasing habitat depth of *C. rupestris*.

| Feature | Expectation |
|---|---|
| **Body size** | Decreases with depth |
| **Body shape** | Becomes more elongated with depth |
| **Mouth gape size** | Increases with depth |
| **Eye size** | Decreases with depth |
| **Gill raker number** | Decreases with depth |
| **Swimbladder mass** | Increases with depth |
| **Liver mass** | Increases with depth |

## Methods

A sample of 139 *C. rupestris* individuals were obtained by trawling in the Northeast Atlantic in August 2018 (see Table 2). Trawls were carried out at five depth levels between 720 and 1830 m. A length-stratified sample of *C. rupestris* was taken from the total obtained at each depth level in order to demonstrate the size range observed. Fish were weighed immediately after capture before being tagged with unique identification numbers and frozen. The frozen specimens were then shipped to Durham where they were stored in freezers. As large quantities of specimens were frozen together, they were partially thawed so that they could be broken apart and frozen again in smaller groups. This meant that specimens were only fully thawed out on the day of their dissection, minimising damage to the organs and body features that might have arisen from repeated thawing and re-freezing.

Table 2: Details of trawls carried out with total catch and sampling of *C. rupestris*.

| **Date** | **Trawl depth** | ***C. rupestris* captured** | ***C. rupestris* included in sample** | **Trawl duration** | **Latitude** | **Longitude** |
|---|---|---|---|---|---|---|
| 27/08/2018 | 720 | 26 | 23 | 30 | 59.20917 | -9.893 |
| 27/08/2018 | 1060 | 208 | 30 | 30 | 59.10583 | -9.8571667 |
| 28/08/2018 | 1830 | 40 | 27 | 30 | 58.69133 | -9.9103333 |
| 28/08/2018 | 1640 | 116 | 30 | 30 | 59.109 | -9.1358333 |
| 29/08/2018 | 1400 | 45 | 29 | 30 | 59.37217 | -8.6388333 |

*Measurement and dissection protocol*

*C. rupestris* specimens were thawed on the day of their dissection, and a set of external measurements between morphometric landmarks were recorded (list of measurements can be found in Table 3). Each of these lengths and distances were measured three times and averaged to maximise accuracy. Distances smaller than 120 mm were measured to the nearest 0.1 mm using dial Vernier callipers. Larger distances were measured to the nearest 1 mm using either large Vernier callipers, spring joint callipers, or a tape measure. Measurements were taken from the left side of all specimens. Due to damage incurred by the specimens during trawling, not all measurements were possible on every individual. The most common form of damage was to the long, fragile tails, many of which were missing or broken. For this reason, total length measurements were not possible for all individuals, and pre-anal fin lengths were used as a substitute overall length measure, as is common in this species (Savvatimsky, 1985; O'Hea et al., 2013).

Table 3: External measurements conducted on *C. rupestris* specimens with their specifications.

| Measure | Specifications |
|---|---|
| **Pre-anal fin length** | Tip of snout to first ray of anal fin along median axis of body |
| **Total length** | Tip of snout to end of tail |
| **Head length** | Tip of snout to posterior edge of operculum |
| **Head depth** | Greatest vertical depth of head anterior to the operculum |
| **Pre-orbital length** | Tip of snout to anterior margin of the orbit |
| **Inter-orbital width** | Distance between left and right orbit as measured from a dorsal view |
| **Orbit diameter** | Distance between anterior and posterior margins of each orbit |
| **Body depth** | Greatest vertical depth of the body posterior to the operculum |
| **Maximum width** | Width of body at its widest point posterior to the operculum |
| **Gape size** | Distance between the jaws when open to their maximum extent |
| **Gape length** | Tip of snout to posterior angle of the mouth along median axis of body |
| **Tail thickness** | Vertical thickness of tail, excluding fin rays, measured 5 cm from the tail tip |

After the external measurements were recorded, dissection began. The left operculum was removed with scissors and the first gill arch removed. The number of gill rakers on the upper and lower limb of the gill arch were then counted. An incision was then made using scissors at the anus, and extended up the ventral midline to the head. A viewing window was cut on the left side of the specimen moving dorsally from this incision. The liver, swimbladder, gonads and heart could then be identified (see Table 4), removed and weighed to the nearest 0.1 g, or for organs weighing less than 1 g, to the nearest 0.001 g. The digestive tract was identified and removed. The pressure decrease experienced when specimens are captured and brought to the surface often causes them to regurgitate their stomachs and further components of the digestive system, so these organs were found to varying degrees either in the main body cavity, or in the everted stomach in the mouth. The pyloric caeca bundle was removed from the digestive tract, and the number of fingers counted. Finally, otoliths were collected from both sides of the skull. With the specimen positioned on its ventral side, an incision was made just behind the bony protrusion of the posterior end of the skull. This incision was made at a 45° angle moving towards the anterior and ventral planes. After this incision was made, it was possible to prise open the top of the skull and remove the otoliths

from their chambers, clean, dry, and store them. All body parts were added to clinical waste bins for incineration after dissection.

Table 4: Internal organs and how they were identified during dissection.

| Organ | Identification |
|---|---|
| **Liver** | Largest organ in the body cavity, with two asymmetrically sized lobes. Beige in colour, appears slightly waxy. |
| **Swimbladder** | Membrane-bound organ attached to the dorsal wall of the body cavity, filled with white, foamy, fatty substance. Membrane occasionally ruptures to release this lipid into the body cavity. |
| **Gonads** | Bifurcate organ found leading to the urogenital ducts, adjacent to the anus. Appearance can vary from small, pink and grainy to swollen and filled with white eggs in non-gravid/gravid females. Male gonads are generally long and thin, containing white or yellow milt. |
| **Heart** | Small red organ found in separate cavity anterior to other organs. |

*Otolith preparation and sectioning*

Otoliths are a good way of estimating age in many fish species, as when sectioned they reveal annuli which can be interpreted as annual growth increments (Lorance et al., 2003). There has been disagreement on whether this technique is appropriate for deep sea species, based on the supposition that the deep sea environment is largely aseasonal. However, there is evidence that the seasonal variations in organic materials transferred from the productive upper layers of the ocean can leave seasonal increments in the otoliths of deep sea fish (Morales-Nin and Panfili, 2005). Several studies have shown that the striations in *C. rupestris* otoliths are laid down annually (Gordon et al., 1995; Gordon and Swan, 1996). Though otolith annuli are difficult to read accurately in this species, especially in larger fish (Lorance et al., 2003), they were collected here to give an additional demographic measure to be analysed only in conjunction with other parameters, such as pre-anal fin length. Both otoliths were removed from each of the 139 specimens dissected. The left otolith from each specimen was then set in epoxy resin to be sectioned and its annuli counted.

Kleer set resin was used to set the otoliths. A 3% hardener to resin ratio was used, and a thin layer was poured into the bottoms of 1.5x4 cm silicone moulds. The moulds were laid flat for the resin to dry. When these resin bases were set, they were removed from the moulds, and a thin black line was drawn down the middle of the bottom of the base. These bases were then restored to the moulds, and well identification numbers (WINs) added to the top right corner of each. Three otoliths were then placed in each mould, with the sulcus, a grove which runs along the proximal otolith surface, facing upwards and positioned perpendicular to the black guidance line marked on the resin base (See figure 2). Fish identification numbers and their corresponding WINs were recorded. More resin was then mixed and added, and the blocks left to dry. These blocks were then sent to colleagues at Marine Scotland Science to be sectioned by a precision sectioning saw along the guidance line marked on the blocks. The annuli were then read by two experienced independent readers on at least two separate occasions. If reader median counts differed by more than 10%, the otoliths were re-read. If discrepancies remained after this point, the otolith in question would be excluded from the sample. An absolute average age was then assigned to each fish.

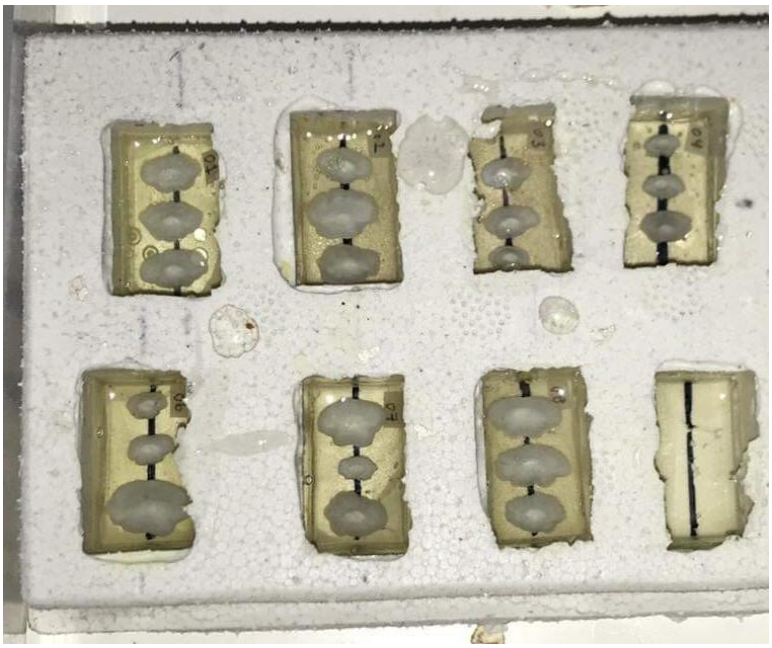

Figure 2: *C. rupestris* otoliths aligned and set in resin blocks. The sectioning plane is cut down the black reference line in the middle of each block.

*Statistical analysis*

All statistical analyses for this project were carried out using IBM SPSS Statistics for Mac version 24.0.

A series of simple, single-variable statistical tests were first carried out for each of the external length and internal mass variables which were measured. Variables were standardised by dividing by either pre-anal fin length (for external length measures) or total mass (for organ masses) to give values which were relative to fish size, removing body size confounds. To minimise the impact of some of the confounds associated with ontogeny, juveniles (categorised as having a pre-anal fin length below 100 mm; N = 17) were removed from these analyses. As there is no great agreement in the literature on size classification for juveniles (see differing categorisations by Gordon and Swan, 1996; Mauchline et al., 1994; Bergstad et al., 2014), this categorisation was chosen because during dissection, it was noted that all individuals with pre-anal fin lengths below 100 mm had undeveloped or indistinguishable gonads. These individuals also appeared to form a distinct group on the lower end of the size range below 100 mm, so were categorised as juveniles.

Normality was assessed at each depth level for every variable using the Shapiro Wilk test. As the ANOVA and Welch ANOVA are robust to Type 1 errors even with slight disruptions to normality (Blanca et al., 2017), normality was deemed acceptable for these tests if at least one depth level showed significant normality according to the Shapiro-Wilk test, and none of the levels which failed this test demonstrated absolute skewness or kurtosis values above 2 or 4, respectively (Kim, 2013). Homogeneity of variances was explored using the Levene statistic. If homogeneity of variances was found and normality was acceptable, an ANOVA with post-hoc Tukey test was used to find differences across the 5 depth levels in the variable being examined. If the assumption of homogeneity of variances was violated while normality at each depth level remained acceptable, a Welch ANOVA was used with a post-hoc Games-

Howell test. If the normality assumption was violated, the non-parametric Kruskal-Wallis H test was used.

The same body-size corrected variables were also tested for changes across wider depth categories, where individuals from shallower and deeper than 1500 m were grouped together to form 'shallow' or 'deep' groups. These broader depth categories were used to test specifically for changes associated with the meso-bathypelagic boundary. The 1500 m cutoff between 'shallow' and 'deep' groups was extrapolated from the recent work by Gaither et al. (2018) who highlighted the relevance of this depth in the genetic segregation which they observed. Normality and homogeneity of variance testing was carried out for both depth levels with the same criteria as described above, and an independent samples t-test was carried out if the data showed acceptable normality. If the data deviated significantly from normality, a Mann-Whitney U test was used.

Gonosomatic index (GSI) was calculated for each individual using the equation:

$$GSI = \frac{gonad\ mass\ (g)}{ungutted\ mass\ (g)} \times 100$$

and reproducing individuals were categorised as having a GSI of ≥ 3 (Bergstad, 1990). A chi-square test was used to test whether there was a relationship between depth level and reproductive status, with Phi and Cramer's V tests to explore the strength of association. The chi-square test was also used to test for relationships between maturity status and depth.

Analysis of weight-length relationships were carried out in accordance with the methods outlined by Neat and Campbell (2013) in order to compare body elongation effectively. For each of the five depth groups, a graph of log transformed mass against pre-anal fin length was created. These graphs allowed the calculation of parameters *a* and *b* as described in the fish length weight relationship equation:

$$W = aL^b$$

where $W$ is total mass (g) and $L$ is length (cm). The parameter $a$ is a coefficient found where the Y-intercept of the graph of log $W$ against log $L$ is log $a$, and $b$ is the allometry coefficient, equal to the slope of the regression line of log $W$ against log $L$ (Schneider et al., 2000). The form factor $a_{3.0}$ was derived for each depth group using the equation:

$$a_{3.0} = 10^{log_{10}(a) - S(b-3)}$$

where $S$ is the slope of the regression line of log $a$ against $b$ (Froese, 2006). $a_{3.0}$ values were then plotted against depth to show any potential changes in body form with depth. A Spearman's rank test was used to search for any correlation between $a_{3.0}$ values and depth.

A simpler analysis of body elongation was also carried out using pre-anal fin length divided by total mass. This new variable was labelled 'elongation' and analysed with a Kruskal-Wallis test across all five depth levels, and an independent t-test either side of the 1500 m boundary.

All of these single trait tests involved repeated implementations of the same tests on different variables in the dataset. Because all variables were measured from the same specimens, there was a risk of increasing chance for Type 1 error when repeating the same tests across all variables (McDonald, 2014). For this reason, the Bonferroni correction was used to reduce critical values and correct for the likelihood of Type 1 error. The critical value of 0.05 was therefore divided by the number of statistical tests of each type conducted to get the corrected critical values. So, for the t-tests, the new critical value became 0.05/12, or $0.0041\dot{6}$. For chi square tests and Mann-Whitney U tests, the new value was 0.05/3, or $0.01\dot{6}$. For standard and Welch ANOVAS, critical value became 0.003125 (0.05/16). For Kruskal-Wallis H tests, the new critical value was 0.05/2, or 0.025.

Multivariate approaches were also attempted. The extremely high level of multicollinearity between variables made tests like the MANOVA or multiple regression inappropriate (Slinker and Glantz, 1985; French et al., 2008). A principal component analysis (PCA) was

therefore attempted in order to group the multicollinear variables into phenotypic components which could then be more effectively analysed. The Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) and Bartlett's test of sphericity were used to confirm that the data were suitable for data reduction, and the PCA carried out. Components with Eigenvalues greater than 1 were then selected for further analysis and analysed using the same methods and tests as described for the single-variable tests above. Bonferroni correction was also implemented for these tests, making the critical value for ANOVAs on the PCA outputs 0.01ȯ (0.05/3), and for t-tests 0.025 (0.05/2).

Dr Thomas Regnier at Marine Scotland Science also carried out a set of clustering analyses using these data. A Random Forest analysis was conducted. This is a computer learning method which uses a large number of decision trees which have been trained using the dataset to group individuals into certain depths according to their morphological characteristics. Each tree is grown using bootstrap draws of around 75% of the data available, with the remaining data being used for cross-validation. Predictions from the 'forest' of trees were then used to predict a class (in this case, depth level) for each individual. Initial Random Forests used all variables, but variables which did not impact class groupings were systematically removed until 60% classification accuracy was reached across all 5 depth groups, and 85% accuracy when the data were split into only two groups shallower and deeper than 1500 m.

An unsupervised approach to the Random Forest analysis was also attempted by Dr Regnier, where the depth of capture of each individual was not used to train the Random Forest, but the method was allowed to make clusters according to morphological characters. These clusters could then be compared to the depth at which individuals were captured to ascertain whether individuals from similar depths were grouped together. Random Forest was used to produce a dissimilarity matrix based on how often two individuals appear in the same terminal node of a tree. Partition Around Medioids (PAM), a clustering algorithm, was then

used to assign each individual to a cluster. Cluster composition at each depth level was then analysed in order to demonstrate whether different depths were dominated by different clusters, suggesting a difference in morphological traits across depths.

## Results

*Demographic and general trends*

A total of 139 individuals were dissected, with estimated ages ranging from 6 to 45 years according to otolith sectioning analysis. All raw data are provided in Appendices 1-3. Several juvenile individuals were identified during dissection from their small size and underdeveloped gonads. However, estimated ages from sectioned otoliths for these small individuals varied substantially. For example, in the 7 individuals with pre-anal fin lengths below 60 mm, the estimated age range was between 6 and 22. Beamish (1979) suggested that accuracy for age estimates in sectioned otoliths is lower in younger individuals, and therefore pre-anal fin lengths were used to divide maturity groups instead of estimated ages. Overall, there is a strong correlation between pre-anal fin length and estimated age (Figure 3; $r^2 = 0.644$, n = 133, $p = 6.199\times10^{-17}$), so pre-anal fin length can be seen as a suitable maturity proxy in this context.

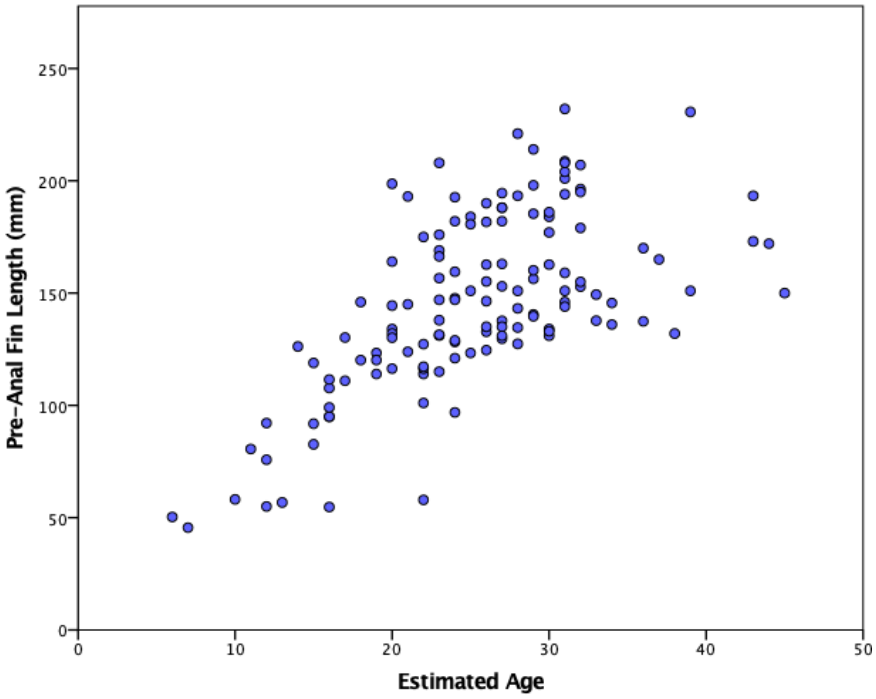


Figure 3: Correlation between pre-anal fin length and age as estimated from otolith annuli.

Juveniles, defined as having a pre-anal fin length below 100 mm, were found at intermediate depths only (see Figure 4a). This association was significant ($\chi^2(4) = 13.759$, $p = 0.008$) with moderate effect size ($\Phi = 0.315$). There was also a significant association between depth and reproductive status. Reproducing individuals, categorised as having a GSI over 3 (Bergstad, 1990), were strongly associated with depths of 1060 m or shallower ($\chi^2(4) = 32.677$, $p = 1\times10^{-6}$; $\Phi = 0.485$; Figure 4b). No significant overall relationship between age and depth was found among the 5 depth groups ($F_{Welch}(4,128) = 1.754$, $p = 0.089$), or across the 1500 m boundary ($t = -1.973$, $p = 0.051$).

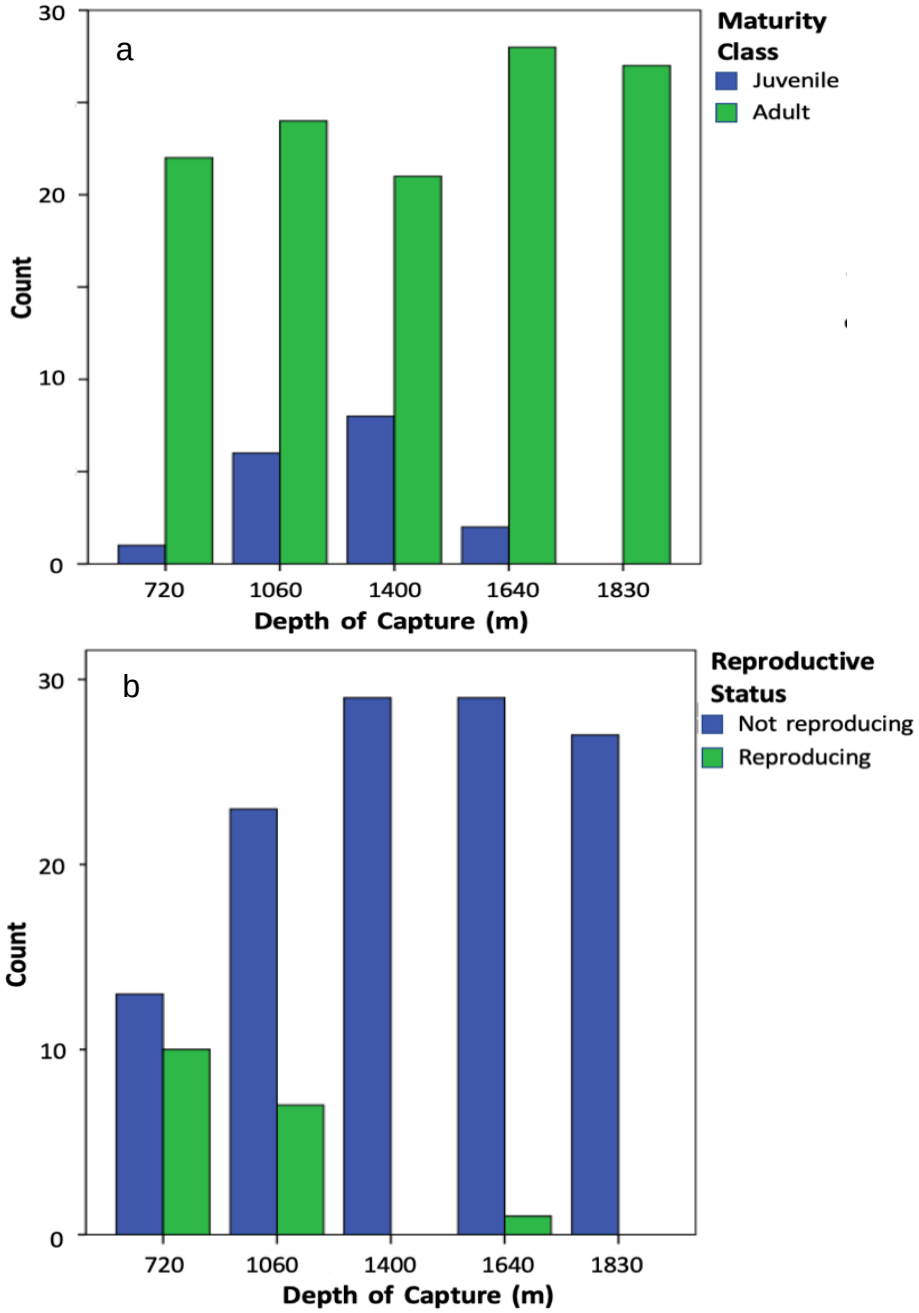


Figure 4: Demographic trends in *C. rupestris* showing a) distribution of juveniles and adults according to depth, and b) distribution of reproducing and non-reproducing individuals.

Pre-anal fin length, here treated as a proxy for total length (Savvatimsky, 1985; O'Hea et al., 2013), decreased significantly in adults as depth increased (Figure 5a; $F_{Welch}(4,117) = 18.035$, $p = 1.a \times 10^{-10}$). Post-hoc Games-Howell testing demonstrated significant differences in pre-anal fin length between nearly all depth groups except those adjacent to one another. The only non-adjacent depth groups with non-significant differences were 1400 and 1830 m ($p = 0.328$), and 710 and 1400 m ($p = 0.009$; Appendix 4). The total mass of adult individuals also demonstrated a significant decrease with increasing depth (Figure 5b; $F_{Welch}(4,117) = 23.590$, $p = 2.521 \times 10^{-12}$). Post-hoc Games-Howell testing revealed significant differences in mass between the two shallowest depth groups and the two deepest depth groups (Appendix 5).

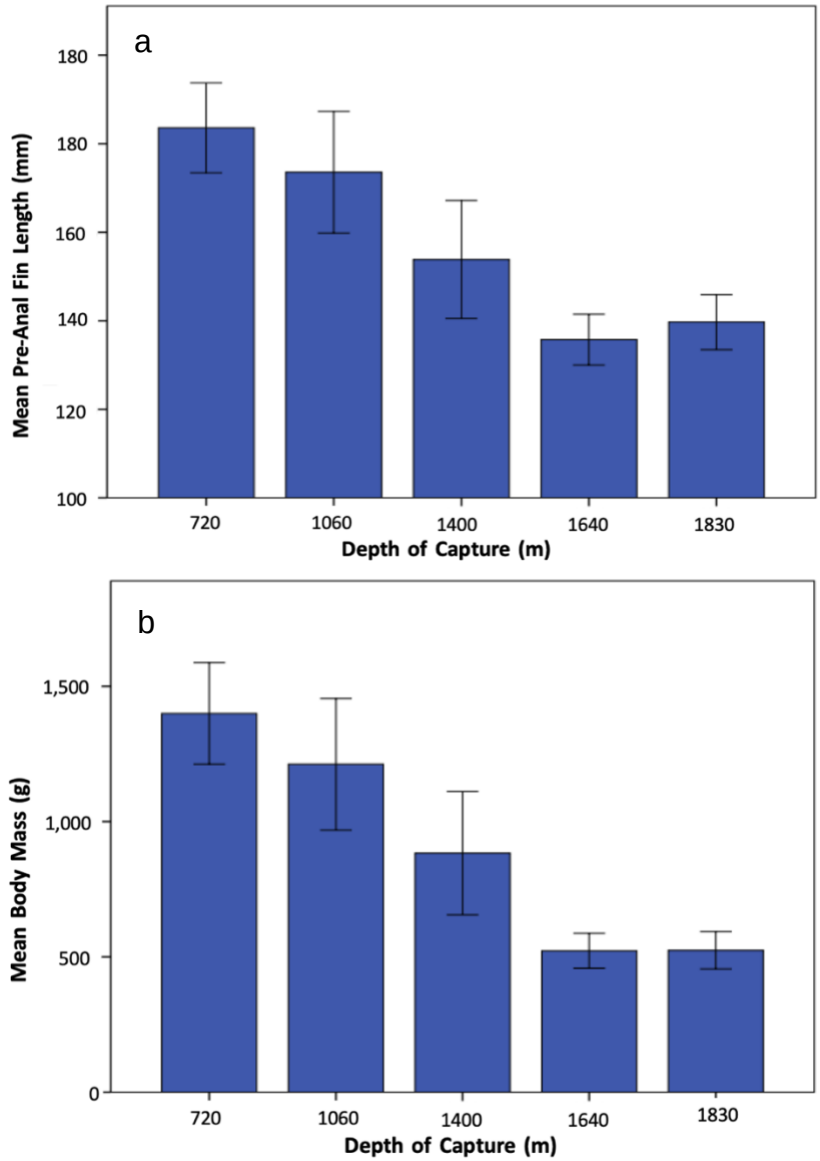


Figure 5: Decreasing body size with habitat depth as demonstrated by a) pre-anal fin length and b) body mass. Error bars are ±2SE.

Body elongation was analysed by plotting form factor $a_{3.0}$ against habitat depth. There was no clear trend in the data and a Spearman's rank test yielded a non-significant result ($r^2$ = -0.300, $p$ = 0.624). When analysed with the 'elongation' measure of pre-anal fin length divided by total mass, however, there was a clear trend for increased elongation at deeper habitat levels (Figure 6). This was confirmed as significant with a Kruskal-Wallis H test across all five depth levels ($\chi^2(4) = 49.229$, $p = 5.231\times10^{-10}$) with a mean rank elongation score of 28.32 for 720 m, 42.83 for 1060 m, 57.86 for 1400 m, 83.11 for 1640 m, and 85.56 for 1830 m. This trend was significant across the 1500 m boundary, with elongation shallower than 1500 m significantly lower (0.183 ± 0.011) than those from deeper habitats (0.070 ± 0.009; $t(120) = -6.601$, $p = 1.169\times10^{-9}$)

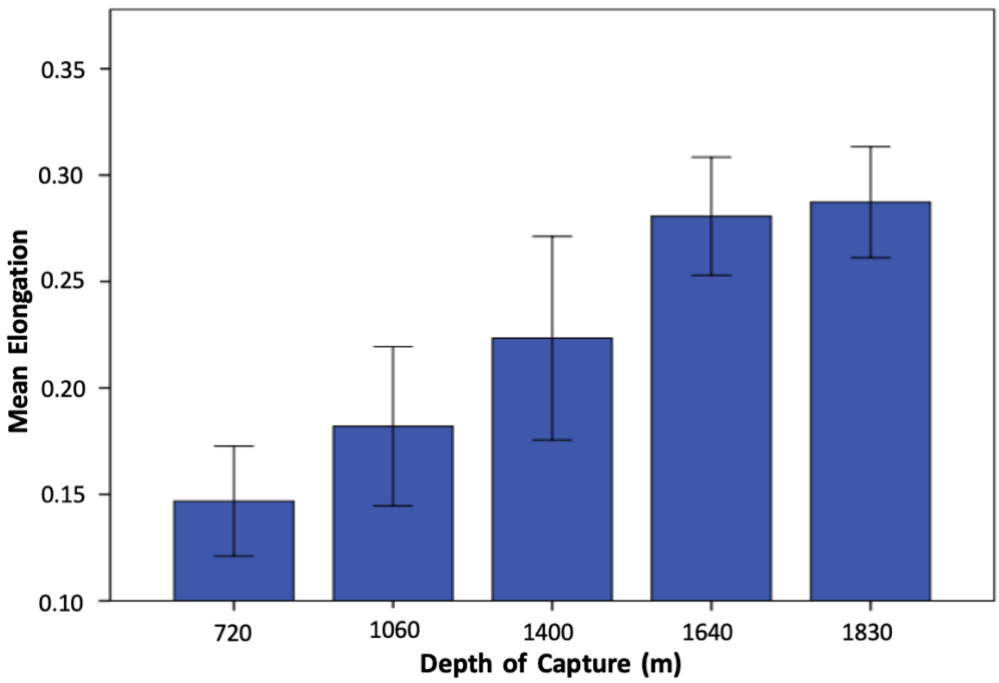


Figure 6: Mean elongation in each depth group. Elongation calculated as pre-anal fin length divided by total mass. Error bars are ±2SE.

*Single trait tests*

In adults, the liver and swimbladder both showed significant increases in relative mass as depth increased, as predicted. For the liver, a significant increase was noted across the five depth levels (Figure 7a; $F_{Welch}(4,117) = 7.724$, $p = 7.281\times10^{-11}$). Post-hoc Games Howell tests showed that this difference was only significant between the 720 m group and the 1640

and 1830 m groups (Appendix 6). When divided into meso- and bathypelagic groups shallower and deeper than 1500 m, however, a clearer result emerged, showing that relative liver mass was significantly lower shallower than 1500 m (0.023 ± 0.015) than deeper (0.034 ± 0.014), $t(120) = -4.059$, $p = 8.8\times10^{-5}$). Swimbladder mass also increased with depth (Figure 7b). This trend is statistically significant across the five depth levels ($F_{Welch}(4,117) = 11.988$, $p = 1.915\times10^{-9}$), with post-hoc Games-Howell tests showing significant differences in the two uppermost depth levels versus the two deepest depth levels (Appendix 7). This trend unsurprisingly remains significant when grouped to shallower and deeper than the 1500 m boundary, with swimbladder masses from shallower than 1500 m significantly lighter than those from deeper than 1500 m ($U = 632.5$, $p = 4.785\times10^{-10}$).

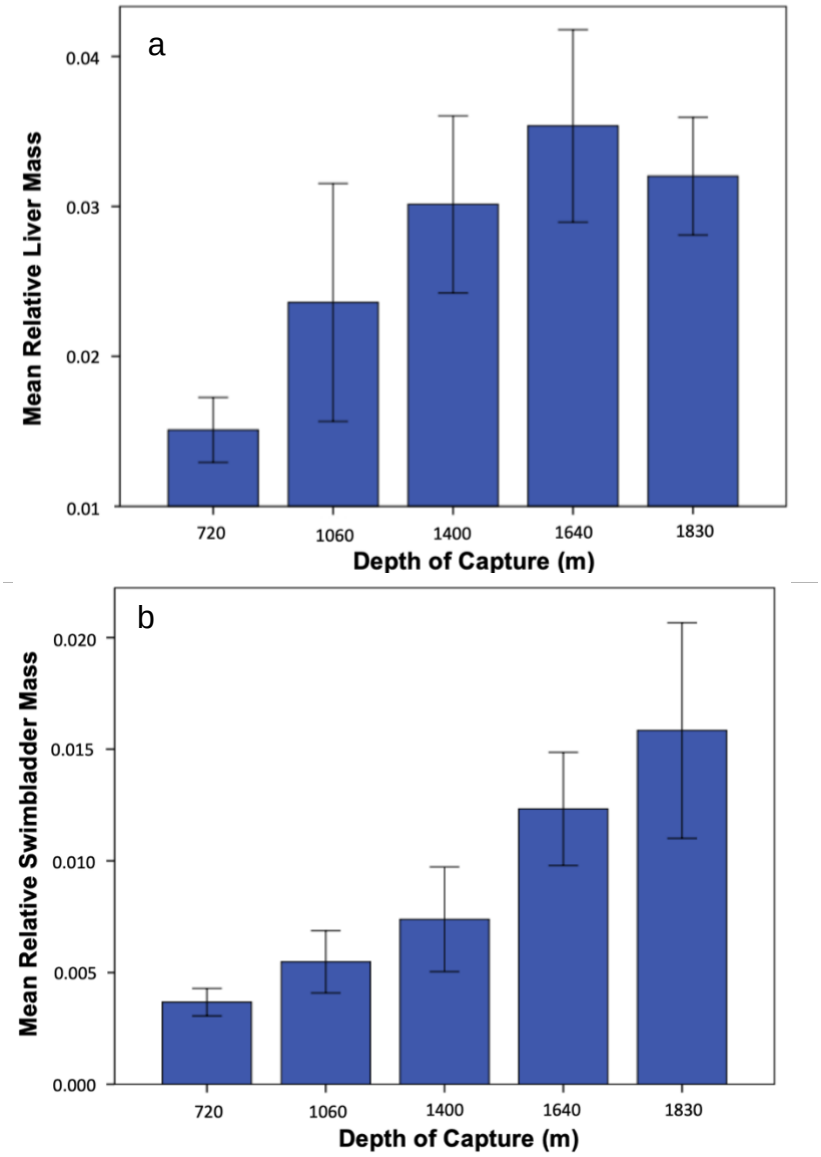


Figure 7: Mean relative a) liver mass and b) swimbladder mass at different habitat depths ±2SE.

Relative heart mass did not demonstrate any trend associated with depth, and a Welch ANOVA did not retrieve a significant result ($F_{Welch}(4,117) = 1.860$, $p = 0.026$). No differences at the 1500 m meso- bathypelagic boundary could be identified either ($t(120) = -0.221$, $p = 0.826$).

Relative maximum body width decreased significantly with depth (Figure 8b; $F(4,117) = 6.581$, $p = 8.2\times10^{-5}$). Post-hoc Tukey testing revealed significant differences only between the 720m depth group and the 1640 and 1830 m groups (Appendix 8). Looking at the 1500 m boundary, maximum body width was shown to be significantly larger in fish from shallower than 1500 m (0.304 ± 0.0312) than deeper (0.282 ± 0.0281; $t(120) = 4.024$, $p = 1.01\times10^{-4}$). No significant trend was observed between relative body depth and habitat depth at the five habitat depth levels ($F(4,117) = 1.611$, $p = 0.176$), or at the two depth levels either side of the 1500 m boundary (0.589 ± 0.0350; $t(120) = 2.205$, $p = 0.029$).

Relative head length showed no obvious trend with depth (Figure 8a), though ANOVA testing across all 5 depth levels yielded a significant result ($F(4,117) = 6.178$, $p = 1.52\times10^{-4}$). Post-hoc Tukey tests demonstrated significantly increased head length in the 1400 m group compared to the 720 and 1060 m groups (Appendix 9). When separated into shallower and deeper than 1500 m groups, no significant difference was found ($t(120)= -1.421$, $p = 0.158$). Relative head depth demonstrated no significant trend at either the five or the two depth group resolutions ($F(4,117) = 2.306$, $p = 0.062$; $t(120) = 1.749$, $p = 0.083$).

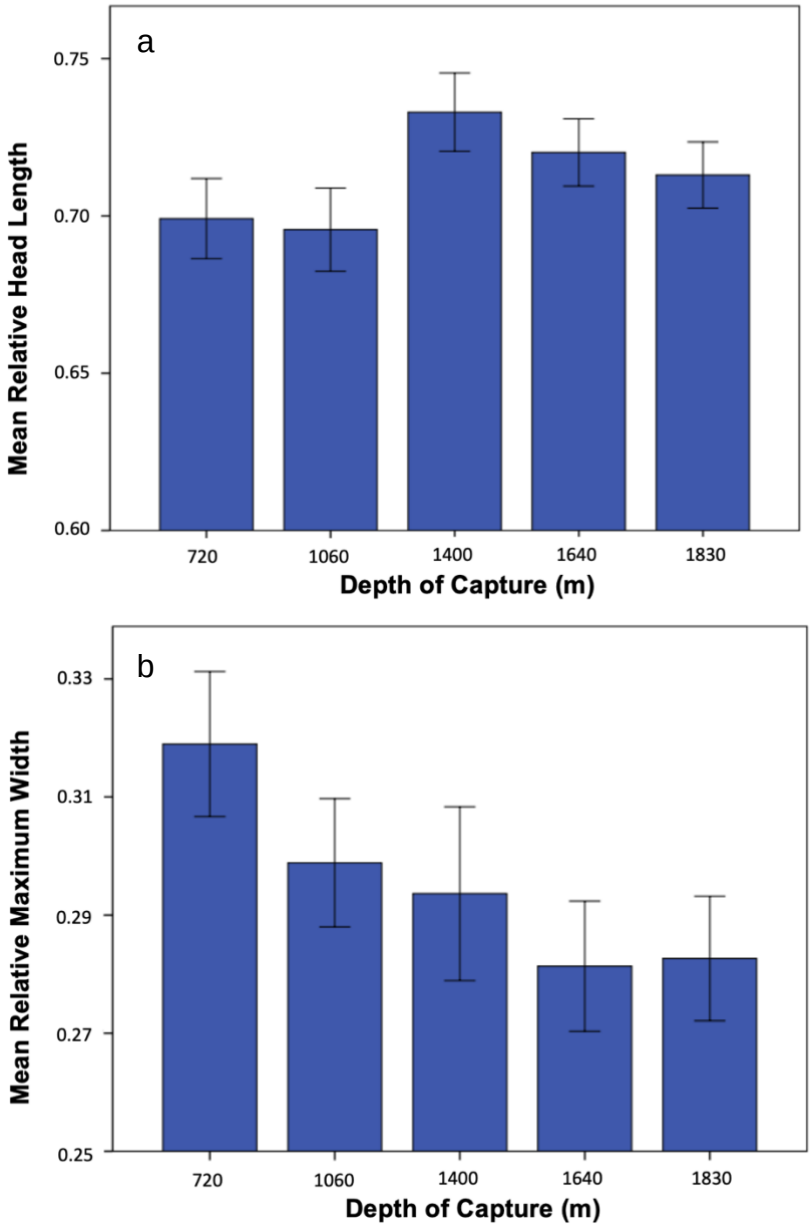


Figure 8: Relative morphological trait sizes at increasing habitat depth: a) head length, b) maximum width. Error bars are ±2SE.

The relative size of the left orbit increased significantly across the five depth levels (Figure 9a; $F_{Welch}(4,117) = 5.585$, $p = 3.36\times10^{-4}$). Post-hoc Games-Howell tests showed significant increases in orbit size in the 1640 m group compared to the 720 m group (Appendix 10). Looking at differences across the 1500 m boundary, orbit size was significantly smaller in fish from shallower than 1500 m (0.214 ± 0.0199) than deeper (0.226 ± 0.0155; $t(119.645) = -3.891$, $p = 1.6\times10^{-4}$). Pre-orbital length showed no significant relationship with habitat depth for any comparison ($F(4,117) = 1.505$, $p = 0.307$; $U = 1745.00$, $p = 0.616$). Relative inter-orbital width yielded a significant ANOVA result ($F(4,117) = 5.474$, $p = 4.48\times10^{-4}$),

though a trend with depth is difficult to discern other than an unusually high mean in the 1400 m group (Figure 9b). Post-hoc Tukey testing revealed the difference in relative inter-orbital width was only significant between the 1060 m group and the unusually high 1400 m group (Appendix 11). When analysed either side of the 1500 m boundary, no significant difference was found ($t(120) = -0.624$, $p = 0.534$).

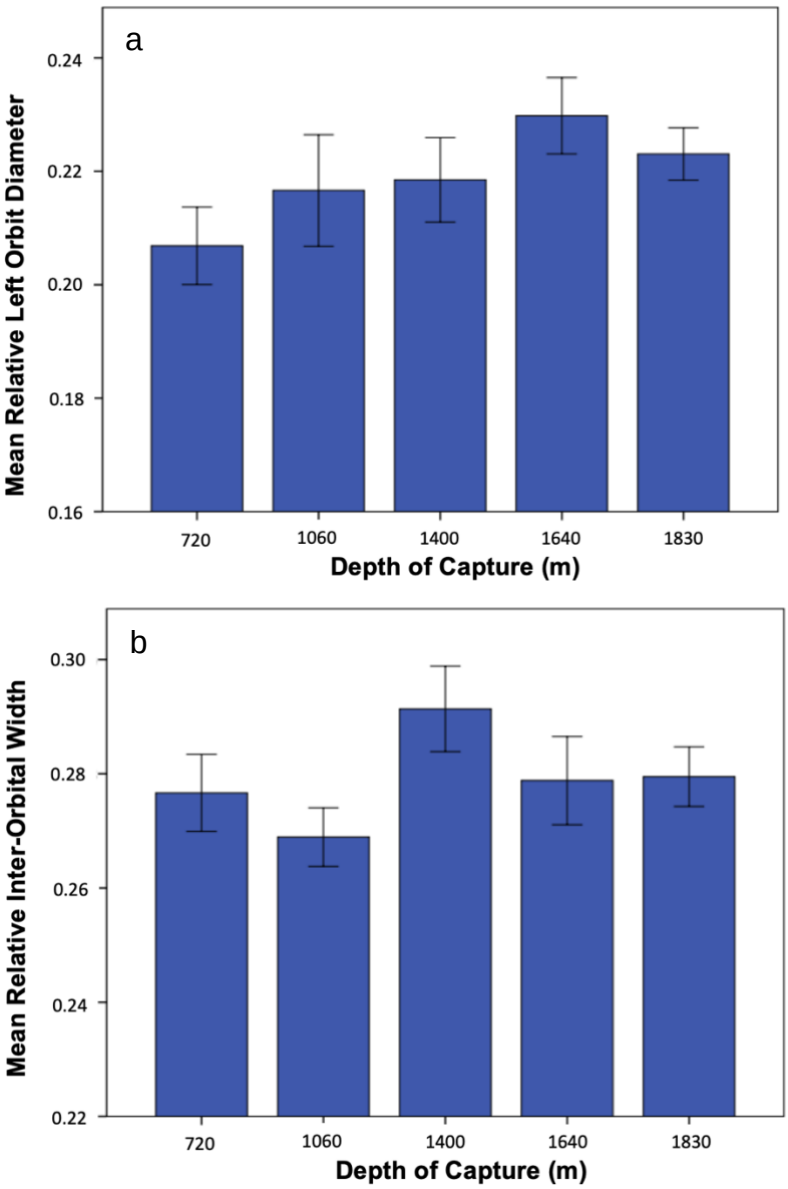


Figure 9: Relative morphological trait sizes at increasing habitat depth: a) left orbit diameter, b) inter-orbital width. Error bars are ±2SE.

Gape size demonstrated a significant increase with habitat depth (Figure 10a; $F_{Welch}(4, 115) = 12.921$, $p = 4.232\times10^{-8}$). Post-hoc Games-Howell tests showed that this difference is significant in the 1830 m depth group compared to both 720 and 1060 (Appendix 12). A

significant difference is just missed between the 1830 and 1400 m depth groups according to the Bonferroni corrected critical values ($p$ = 0.003144). A significant difference between the 720 and 1640 m groups is also observed, demonstrating a general difference in gape size between the shallowest and deepest depth groups. This finding is echoed when gape size is explored above and below the 1500 m boundary: individuals living shallower than 1500 m have significantly smaller relative gape sizes (0.392 ± 0.050) than those living deeper than 1500 m (0.436 ± 0.028; $t(118) = -5.86$, $p = 4.24\times10^{-8}$). Across the five depth levels, gape length does not demonstrate a very clear trend (Figure 10b), although ANOVA testing showed significant relationship between gape length and depth ($F(4,116) = 9.660$, $p = 8.868\times10^{-7}$). Post-hoc Tukey testing revealed that gape length is significantly higher at 1640 m compared to both 1400 and 1830 m (Appendix 13). A clearer trend is observed when only two depth levels are considered either side of 1500 m: individuals living shallower than 1500 m have significantly smaller relative gape lengths (0.236 ± 0.018) than those living deeper than 1500 m (0.246 ± 0.022; $t(119) = -3.09$, $p = 0.002$).

Gill raker number showed no obvious trend across depth levels. This was confirmed by a non-significant Kruskal-Wallis H test across the five depth groups ($\chi^2(4) = 8.157$, $p = 0.086$), and a non-significant Mann-Whitney U test across the two depth groups separated at 1500 m ($U = 1683.5$, $p = 0.784$). Pyloric caeca finger count demonstrated no significant trend across either all five depth levels ($F(2,112) = 1.65$, $p = 0.167$), or across the 1500 m boundary (26.08 ± 3.597; $t(115) = 2.192$, $p = 0.03$) with the Bonferroni-adjusted critical value of $0.0041\dot{6}$.

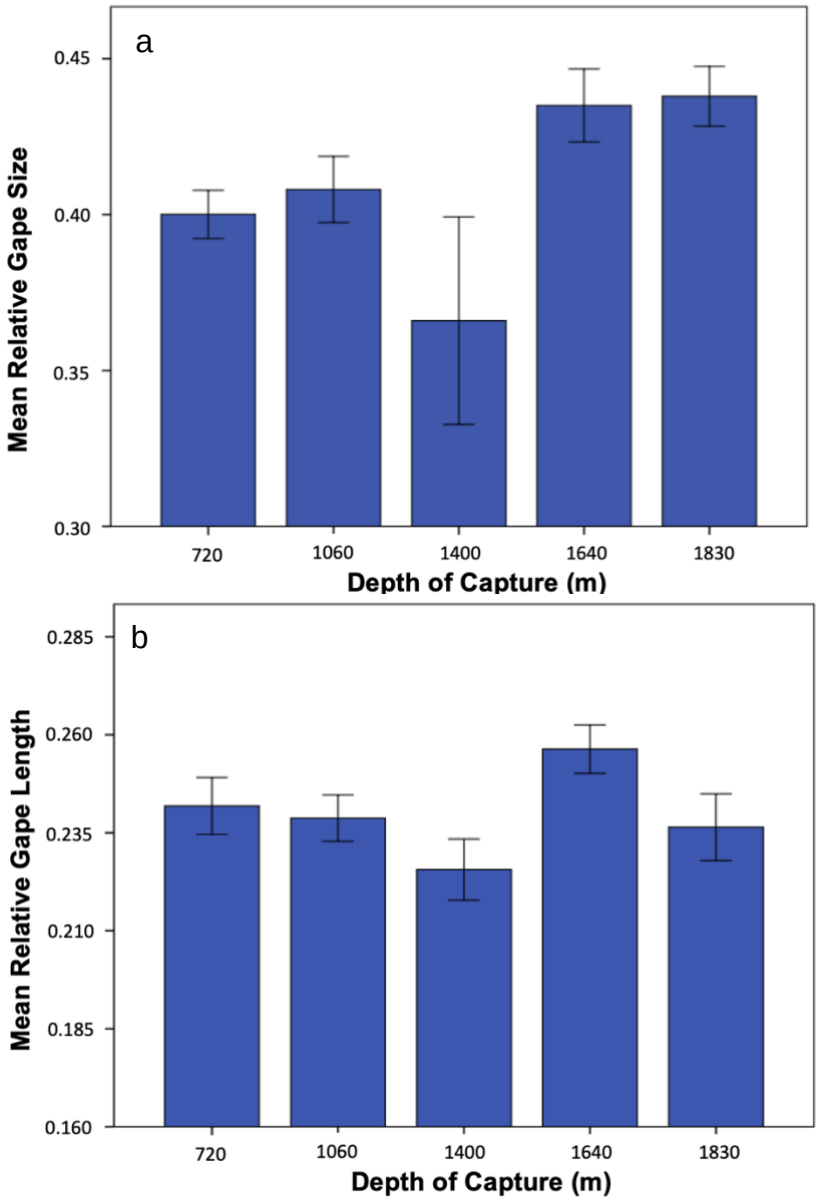


Figure 10: Relative morphological trait sizes at increasing habitat depth: a) gape size, b) gape length. Error bars are ±2SE.

*Principal component analysis*

The principal component analysis was carried out using only adults (pre-anal fin length above 100 mm). 14 morphometric variables were included. They were: pre-anal fin length, head length, head depth, pre-orbital length, inter-orbital width, left orbit diameter, body depth, maximum width, gape size, gape length, liver mass, swimbladder mass, heart mass, and total mass. These data were deemed highly suitable for data reduction using Bartlett's test of sphericity ($p$ = 0) and the KMO (KMO = 0.939).

The first factor to be extracted was extremely robust, with a high eigenvalue of 10.851, and it accounted for a huge 77.5% of total variance in the data. The second factor had an eigenvalue of 1.033, and accounted for a further 7.38% of variation. The third factor extracted had an eigenvalue of only 0.748, and accounted for a further 5.35% of variation. This factor was selected despite its low eigenvalue for its surprising and unique component properties, which will be discussed later. Taken together, these three components account for 90.2% of total variation in this dataset.

Table 5: Component matrix from principal component analysis.

| **Variable** | **Component** | | |
|---|---|---|---|
| | 1: Size | 2: Lipid Storage | 3: Energy Allocation |
| **Pre-anal fin length (mm)** | .981 | -.012 | .061 |
| **Head length (mm)** | .985 | -.022 | .020 |
| **Head depth (mm)** | .976 | -.038 | .029 |
| **Pre-orbital length (mm)** | .904 | -.008 | .020 |
| **Inter-orbital width (mm)** | .966 | .010 | .002 |
| **Left orbit diameter (mm)** | .934 | -.080 | .080 |
| **Body depth (mm)** | .973 | -.045 | .022 |
| **Maximum width (mm)** | .947 | .039 | -.090 |
| **Gape size (mm)** | .774 | -.402 | .044 |
| **Gape length (mm)** | .934 | -.110 | .067 |
| **Liver mass (g)** | .474 | .631 | -.561 |
| **Swimbladder lipid mass (g)** | .425 | .669 | .592 |
| **Heart mass (g)** | .817 | .022 | -.236 |
| **Mass** | .978 | -.040 | -.029 |

Looking at the component matrix (Table 5), it was possible to see that the first factor was correlated strongly with all the variables, with the exception of liver mass and swimbladder mass, which showed lesser correlations. It was clear that this component was a simple combination of all the size relationships between the variables, and was therefore labelled ‘Size’. The second factor was correlated strongly with swimbladder mass and liver mass. All other variables showed minimal negative correlations to this component, with the exception of gape size, which showed an intermediate negative correlation. This component therefore appeared to be related to lipid storage in the liver and swimbladder, where increased lipid storage is related to minor reductions in overall body size. This component was labelled

‘Lipid Storage’. The third component was intriguing, as it demonstrated a strong negative correlation with liver mass, but an equal positive correlation with swimbladder mass. All other variables had minimal positive correlations to this component, except heart mass, which had a small negative correlation. That liver mass and swimbladder mass could be associated positively in the Lipid Storage component, but inversely in this component, was surprising. This suggested that there was some kind of energy allocation trade-off between the lipids in the swimbladder and the liver, where at the higher scores of this component, more energy is allocated to storage in the swimbladder relative to the liver. This component was therefore labelled ‘Energy Allocation’, and included in analyses in order to gain a more complete picture of how energy storage and allocation might change according to these contrasting components with depth.

With the components established and named, they could then be tested for associations with depth. Size was analysed across all five depth groups and was shown to decrease significantly with depth (Figure 11a; $F_{Welch}(4,115) = 16.864$, $p = 7.18 \times 10^{-10}$). Post-hoc Games-Howell testing revealed significant decreases in the Size component in the two deepest depth groups compared to the two shallowest, as well as a significant decrease in the 1400 m group compared to the 720 m group (Appendix 14). This trend was also significant across the 1500 m boundary, where the mean Size in fish from shallower than 1500 m was significantly higher (0.481 ± 1.043) than that in those from deeper than 1500 m (-0.588 ± 0.519; $t(99.151) = 7.295$, $p = 7.47\times10^{-11}$). Lipid Storage increased as depth increased (Figure 11b). This was confirmed with a significant result from a Welch ANOVA ($F_{Welch}(4,115) = 8.346$, $p = 5.447 \times10^{-7}$). Post-hoc Games-Howell testing revealed a significantly lower value in Lipid Storage in the 720 m group compared to all other depth groups (Appendix 15). Increased lipid storage at increased habitat depths was also demonstrated either side of the 1500 m boundary with a Mann-Whitney U test ($U = 1113$, $p = 4.14\times10^{-4}$). Finally, while Energy Allocation increased with increasing depth (Figure 11c), an ANOVA yielded a

narrowly non-significant result across the five depth levels with the Bonferroni-adjusted critical value of $0.01\dot{6}$ ($F(4,115) = 2.676$, $p = 0.035$). When means were compared across the 1500 m boundary, however, an independent t-test demonstrated that Energy Allocation was significantly lower shallower than 1500 m (-0.240 ± 1.024) than deeper (0.293 ± 0.895; $t(118) = -2.998$, $p = 0.003$).

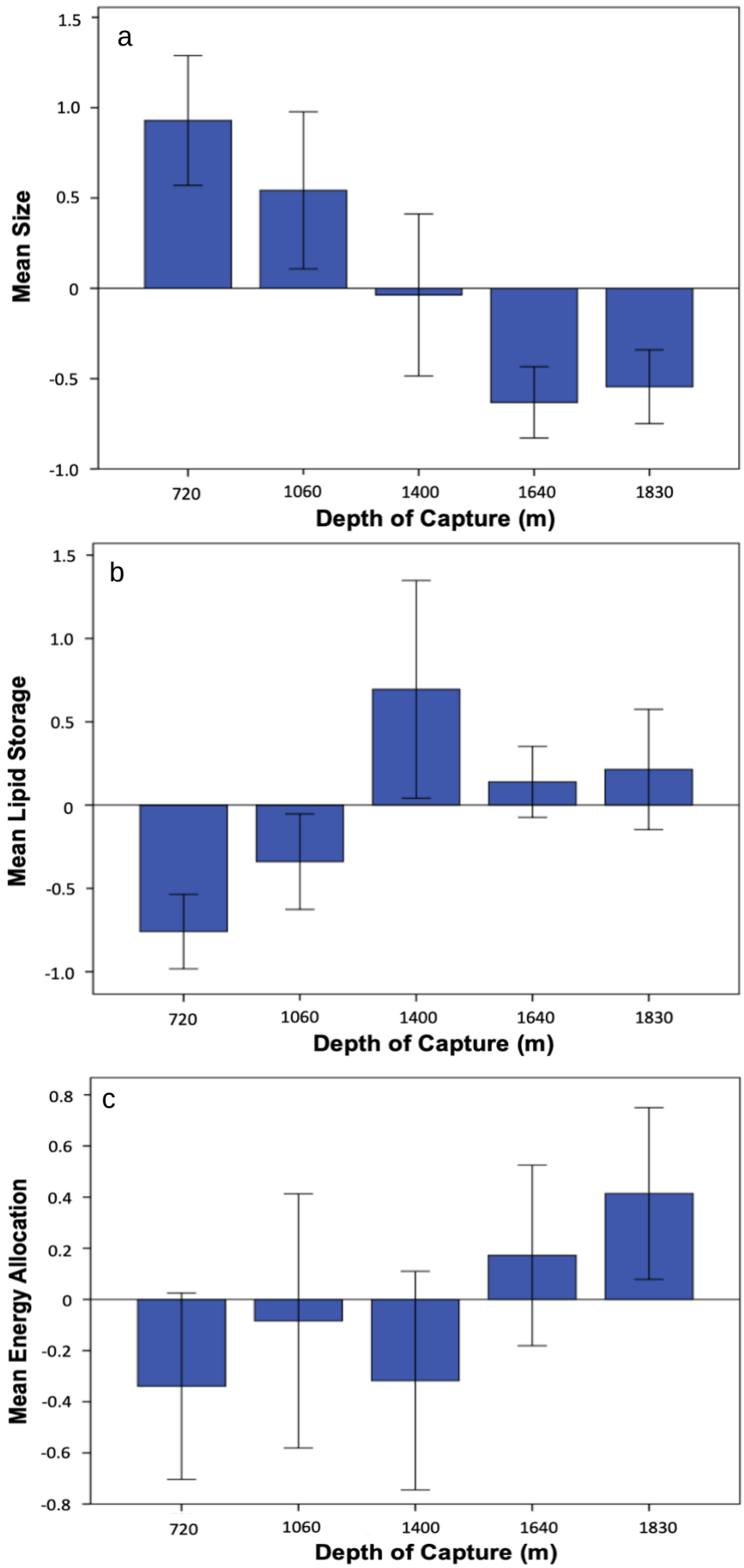


Figure 11: Components from PCA at increasing habitat depths. a) Size, b) Lipid Storage, c) Energy Allocation. Error bars are ±2SE.

*Clustering analyses*

In the first round of Random Forest, only 56.1% of individuals were classified to the correct depth (Table 6). Relative importance for each variable was therefore calculated using mean decrease in accuracy, referring to the decrease in classification accuracy when a variable is not included, and Gini index scores (Table 7). Swimbladder mass and gape length were noted here as having particularly high relative importance for depth classification. Any variable with a mean decrease in accuracy below 0.002 was excluded, allowing 60% classification accuracy to be reached across all five depth groups. Classification accuracy reached 85% when depth levels were grouped together into 'shallow' and 'deep', referring to depths shallower or deeper than the 1500m boundary, respectively (Table 8).

Table 6: Grouping outcome of the first round of Random Forest analysis. 56.1% grouping accuracy.

| **Real depth** | **Predicted depth** | | | | | **Class error** |
|---|---|---|---|---|---|---|
| | 720 | 1060 | 1400 | 1640 | 1830 | |
| **720** | **15** | 4 | 2 | 1 | 0 | 0.32 |
| **1060** | 6 | **9** | 3 | 6 | 1 | 0.64 |
| **1400** | 2 | 3 | **16** | 0 | 2 | 0.30 |
| **1640** | 2 | 5 | 1 | **17** | 3 | 0.39 |
| **1830** | 1 | 1 | 0 | 6 | **17** | 0.32 |

Table 7: Relative classification importance for each variable included in Random Forest analysis. Values for swimbladder mass and gape length are highlighted as particularly important for classification.

| Variable | 720 | 1060 | 1400 | 1640 | 1830 | Mean Decrease Accuracy | Mean Decrease Gini Index |
|---|---|---|---|---|---|---|---|
| **Tail phentype** | 0.014 | 0.002 | 0.002 | 0.000 | 0.001 | 0.003 | 0.994 |
| **Head length** | 0.007 | 0.018 | 0.057 | 0.000 | 0.011 | 0.018 | 7.803 |
| **Head depth** | 0.000 | 0.008 | 0.022 | 0.014 | -0.004 | 0.008 | 6.929 |
| **Pre-orbital length** | 0.015 | 0.005 | -0.001 | -0.004 | 0.002 | 0.002 | 6.343 |
| **Inter-orbital width** | 0.009 | 0.030 | 0.033 | 0.002 | -0.011 | 0.012 | 8.950 |
| **Orbit diameter** | -0.001 | 0.000 | 0.001 | 0.029 | 0.013 | 0.009 | 6.751 |
| **Gape size** | 0.009 | 0.009 | 0.032 | 0.012 | 0.023 | 0.016 | 9.313 |
| **Gape length** | 0.035 | 0.011 | 0.104 | 0.071 | 0.010 | **0.046** | **12.472** |
| **Gill raker count** | 0.003 | 0.002 | 0.013 | 0.004 | 0.005 | 0.006 | 4.107 |
| **Liver mass** | 0.044 | 0.022 | -0.005 | 0.018 | 0.013 | 0.018 | 8.997 |
| **Swimbladder mass** | 0.108 | 0.023 | 0.049 | 0.062 | 0.130 | **0.071** | **16.193** |
| **Heart mass** | 0.011 | 0.000 | -0.001 | 0.010 | 0.063 | 0.016 | 8.583 |

Table 8: Classification outcome when depth levels are grouped into ‘shallow’ (shallower than 1500m) and ‘deep’ (deeper than 1500m). 85% classification accuracy.

| Real depth | Predicted depth | | Class error |
|---|---|---|---|
| | Shallow | Deep | |
| **720** | 22 | 0 | 0.00 |
| **1060** | 20 | 5 | 0.20 |
| **1400** | 21 | 2 | 0.09 |
| **1640** | 7 | 21 | 0.25 |
| **1830** | 4 | 21 | 0.16 |

The unsupervised learning technique yielded the best results by creating three distinct morphological clusters. The contribution each of these clusters made to each depth level was then analysed (Figure 12). Cluster 1 is a key contributor to the 1640 and 1830 m depth levels, but is present in shallower depths. Cluster 2 is more prevalent at greater depths but also present at shallower levels. Cluster 3 is predominantly associated with shallow depths and is almost absent from the two deepest habitat levels.

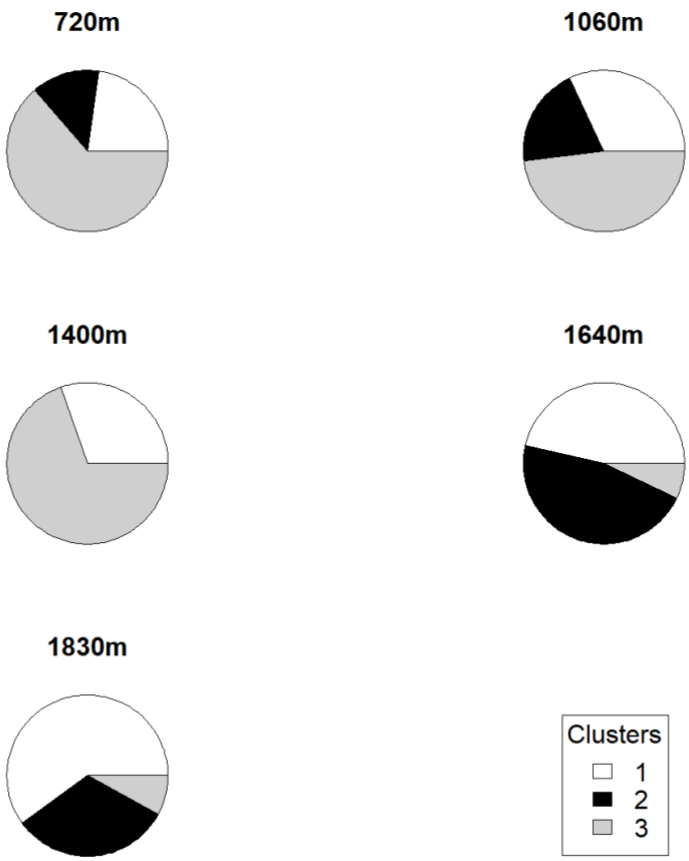


Figure 12: Graphical representation of the cluster composition for the different depths from unsupervised Random Forest analysis.

## Discussion

*Demographic and general trends*

That juveniles were found mainly at intermediate depths, peaking at 1400 m, was somewhat surprising, as the literature suggests that juveniles tend to occupy much shallower habitats. Gaither et al. (2018) found the majority of juveniles shallower than 1200 m, and Gordon and Swan (1996) showed that juveniles were most abundant at 1000 m. This discrepancy may be partially due to classification dissimilarities. For example, Gordon and Swan (1996) analysed only individuals with head lengths below 50 mm, classing them as juveniles, while in another study of juvenile *C. rupestris*, they were classified as individuals with total lengths of 80-100 mm (Mauchline et al., 1994). In yet another study, juveniles are classed as having pre-anal fin lengths ≤ 50 mm (Bergstad et al., 2014). These inconsistencies show there is no

standard size classification for juveniles in this species. The classification used here, where juveniles have pre-anal fin lengths of ≤ 100 mm, does suggest that the individuals marked as juveniles here were larger than juveniles in other studies. This could be the result of a relatively large mesh size being used for these trawls (20 mm in this study compared to the 6 mm mesh used by Bergstad et al., 2014). Longmore et al. (2011) demonstrated that while egg, larval and juvenile life stages are restricted to shallow habitats, individuals migrate deeper with maturity. This could explain why the slightly larger, and therefore older juveniles of this study were found at intermediate depths rather than the shallowest depths.

The distribution of reproducing individuals was more in line with expectations from the literature. The strong preference for shallower waters for these individuals supports other reports of mesopelagic spawning in this species (Bergstad et al., 2014). It is therefore possible that the strong association of reproducing individuals with shallower habitats is a result of reproductive vertical migration, a process which Savvatimsky (1982) proposed after observing seasonal changes in sex ratios at particular depths. If the reproducing individuals had indeed originated from a variety of different depths, this might create a confound in the analysis of morphological data, as individuals who had settled at deeper levels according to their morphological characteristics, but had moved higher to reproduce, would be counted as living permanently at that shallower level. The incomplete data on reproductive migration, however, coupled with the relatively small number of reproducing individuals collected (N=18), mean that reproductive migration cannot be fully established as a confound.

*Body size*

There is clearly some level of morphological segregation occurring across the depth gradient inhabited by *C. rupestris*. Perhaps the most striking change which occurs with increasing depth is the overall reduction in body size. Multiple measures of body size: pre-anal fin length, body mass, and the PCA extracted Size component, all demonstrate significant decreases as depth increases. The reduction in these body size metrics was the most

significant when compared between the two shallowest and two deepest depth groups, suggesting that, as proposed using genetic evidence by Gaither et al. (2018), some key change is occurring across the meso- bathypelagic transition. However, it must be established whether the decreasing body size observed with increasing habitat depth is a result of adaptation, and therefore related to the genetic segregation observed by Gaither et al. (2018), or simply the result of individuals living at deeper levels having less access to food resources and therefore having less opportunity to grow large.

With juveniles excluded from body size analyses, and no significant relationship between fish age and habitat depth, it is certain that the smaller body sizes observed at deeper levels is not a result of younger individuals populating deeper waters at higher frequencies. In fact, there is evidence that migration to deeper habitats is carried out primarily by adults (Longmore et al., 2011, Gaither et al., 2018). In addition, Bergstad (1990) showed using Bertalanffy growth curves that growth in both males and females becomes asymptotic at around 20 years of age. The small proportion of adults under 20 years old (N=10 shallower than 1500 m, N=8 deeper than 1500 m) suggests that the trend observed cannot be the result of decreased growth in bathypelagic individuals due to reduced food resources. These lines of evidence suggest that deeper habitats are not the cause of the smaller body sizes observed, but rather that smaller fish migrate selectively to those depths as they mature. This is in accordance with Thiel's size-structure hypothesis (1975; 1979), which links reduced body size at increasing habitat depths to reduced food requirements, making populations of smaller individuals more suitable at greater depths where food is limited (Ramirez-Llodra et al., 2010). Most examples of Thiel's size-structure hypothesis use multiple species and clades (e.g. Rex et al., 2006), so finding such a pattern within a single species is highly notable.

Thiel's size-structure hypothesis is often linked to reproductive potential, where the advantage of small body sizes in deeper habitats is that more individuals are supported,

allowing a sufficient population size for effective reproduction (Kaariainen and Bett, 2006; Ramirez-Llodra et al., 2010). This is unlikely to be the case in *C. rupestris* however, since the genetic evidence for non-assortative mating (Gaither et al., 2018) combined with the potential for vertical migrations to shallower waters for reproduction (Savvatimsky, 1982; Bergstaad, 2013) suggest that the deeper fish are not reproductively isolated from their shallow living conspecifics. This means there is no need to maintain a population size sufficient for reproduction at increased habitat depths. The reproductive implications of the size-structure hypothesis therefore do not apply in this case, but the key tenets are still valid: smaller body size means lower food requirements, which are an advantage for living in deeper habitats.

Overall, the evidence supports Gaither et al.'s (2018) proposal that individuals with specific genetic markers migrate to settle at depths determined by their genotype. In this case, there must be some set of genetic markers which result in smaller body sizes in adult fish, and those fish must migrate to deeper habitats as they mature. Deeper habitats are more suitable to the smaller individuals as their lesser food requirements mean that they are able to sustain themselves on the limited food resources of the bathypelagic.

*Body shape*

Body elongation in deeper habitats was hypothesised for this species following evidence from Neat and Campbell (2013), who linked a trend for more elongate, slender bodies in deeper living fish species to improved efficiency of anguilliform swimming modes. Neat and Campbell used the form factor $a_{3.0}$ to demonstrate elongation, and when applied to our data, there was no trend in $a_{3.0}$ in response to habitat depth. There are several potential methodological reasons for this lack of a pattern, however. Firstly, this is a method which uses parameters yielded by multiple regression lines, and was used by Neat and Campbell to analyse huge datasets including 98 species. It is possible that our dataset was simply not large enough to yield accurate regression lines and numbers for the parameters. This was

particularly true for the parameter *S*, which was necessary for the final calculation of $a_{3.0}$ values, and which had to be calculated as the slope of a regression line with only five data points, each representing one depth group. Furthermore, length-weight relationships and calculations of $a_{3.0}$ were designed as a function of the total length of fish, not the pre-anal length as used here. It is possible that any level of body elongation that would affect $a_{3.0}$ in fish from deeper habitats was found in the tail, meaning that it could not be properly measured and so was not expressed in these analyses. Methodological limitations in $a_{3.0}$ analysis explain why, when body elongation was analysed in a simpler way using pre-anal fin length divided by mass, a much clearer trend emerged. This separate analysis suggests that body mass does in fact decrease relative to length with increasing habitat depth in *C. rupestris*, suggesting a level of elongation comparable to that shown by Neat and Campbell (2013). This significant result should not be overlooked, and it can be accepted that body elongation, facilitating more energy efficient anguilliform swimming in the bathypelagic, might occur in *C. rupestris*, even if not demonstrated by $a_{3.0}$ values, which are more tailored to large phylogenetic studies.

Besides form factor analysis and direct measures of elongation, other metrics of body shape implied a tendency for more slender morphologies at deeper levels. The significant decrease in body width as habitat depth increased could be linked to improved streamlining to reduce drag and improve swimming efficiency. A study on fluid mechanics, however, suggested that narrow body widths are actually less energy efficient for anguilliform swimming, and that any advantage incurred from smaller body widths would be in improved speed (van Rees et al., 2013). This is at odds with the assumption that swimming efficiency is being selected for in deeper habitats to make up for the low food availability. It is possible therefore that decreased body width is not an adaptation to the bathypelagic zone, but rather a response to the poor food availability in this region. The thickest part of the body, where the maximum width measurement was taken, was invariably in the trunk, where the main organs were

found. It is possible that lesser access to food in the deeper habitats resulted in fish who were captured with less food passing through the digestive system, and so had smaller maximum body widths. Alternatively, smaller body widths in deeper habitats may simply be another manifestation of the selection for smaller overall body sizes at deeper levels as discussed above.

One surprising body shape finding was that eye size, expected to decrease with depth to improve energy efficiency in the dark bathypelagic, actually increased. That this finding was significant across the meso-bathypelagic boundary was particularly striking, as the general trend in the literature is for reduction in eye size in the bathypelagic compared to the mesopelagic (Warrant et al., 2003). Light levels are at zero in the bathypelagic, with the only sources of light being small bioluminescent organisms. According to Warrant (2000), increasing eye size is ineffective in the bathypelagic as it merely improves long-distance sensitivity to bioluminescent point signals, which he suggests are out of range for deep sea fish, who are "weak, and must swim very slowly". This is a broad generalisation to make about all deep sea fishes however, and the elongated body form of *C. rupestris* actually suggests it is an effective anguilliform swimmer (Neat and Campbell, 2013). Though it has been suggested that *C. rupestris* does not have the capacity to perform long-distance swimming migrations (Lorance et al., 2001), behavioural studies have shown them engaging in active swimming behaviours (Lorance and Trenkel, 2006; Neat, 2017). Studies of other grenadiers with similar body shapes and swimming modes also support *C. rupestris* as active swimmers. *C. yaquinae* and *C. armatus* demonstrated active foraging behaviours when monitored with acoustic tagging (Priede et al., 1990), with the latter species demonstrating swimming speeds of almost 0.7 $ms^{-1}$ in observations from a free-fall camera vehicle (Priede et al., 1994). These studies all indicate that *C. rupestris* are capable of swimming beyond the "few tens of metres" that Warrant (2000) says would become detectable with larger eyes. It can therefore be suggested that the trend for larger eyes in bathypelagic *C. rupestris* is linked

to improved long-range detection of bioluminescent prey items: an important advantage in an environment where food is scarce.

The morphology of the mouth and gill raker structure were expected to change with depth, with a supposed trend for larger mouth gape sizes and larger gaps between the gill rakers in bathypelagic fish, allowing them to access a wider range of food resources. While no significant trend was observed in gill raker number in the first gill arch (which corresponds to gap size between gill rakers), strong patterns were observed in mouth size. Gape length and gape size both increased significantly relatively to body size with increasing depth, and the change was particularly apparent when compared across the meso-bathypelagic boundary. Further to this, gape length was identified as having a key role in improving classification accuracy in the supervised Random Forest analysis. This strong evidence for selection for bigger mouth sizes in the bathypelagic can be linked simply to the ability to access a wider range of food resources in an environment where food is limited (Drazen and Sutton, 2017). Diets can be more specialised in the mesopelagic, where there is a greater abundance of food, and this is reflected in the relatively smaller gape sizes and lengths observed.

*Energy storage*

Energy storage in both the liver and swimbladder followed the expected trends. Relative liver mass increased significantly with increasing habitat depth, and the difference was shown clearly across the meso-bathypelagic boundary. Large lipid energy stores, which in *C. rupestris* are found mainly in the liver (Drazen, 2002), signify long periods of fasting in deep sea teleosts (Musick and Cotton, 2015). The food limitations of the bathypelagic mean that *C. rupestris* individuals living within this zone might go long periods without a meal, so appropriate energy storage in the liver is crucial. Further exploration of this trend is required to determine whether this is a plastic response to poor food availability, or whether genetic markers such as those identified by Gaither et al. (2018) create a tendency for high

energy storage in the liver in certain individuals, who then migrate to the bathypelagic as this zone is more suitable to their morphology. Experimental studies in other teleosts have suggested that energy storage in the liver is essential in fish species adapted to periods of food shortage, and that liver energy stores are the first to be replenished after feeding (Collins and Anderson, 1995; Rios et al., 2006). Plasticity in liver storage is no doubt crucial, and effective adaptations must be in place to facilitate the formation of such effective energy reserves. It therefore remains to be seen whether the adaptation for large energy reserves in the liver in periods of food deprivation is present across all *C. rupestris* individuals, or only those who live in the food limited bathypelagic.

Swimbladder lipid mass increased as expected with increasing habitat depth. The strongly significant difference in relative mass between the uppermost and lowermost depth levels, combined with the identification of this variable as important in depth classification in the Random Forest analysis, demonstrate the strength of this trend. There are three key reasons that lipid mass in the swimbladder might increase with habitat depth. First, lipids might facilitate oxygen offloading for swimbladder inflation in environments deeper than around 1500 m, where it is possible that the Root effect becomes less effective (Scholander, 1954; Phleger, 1998). Second, increased swimbladder lipids might protect against the diffusional loss of swimbladder gases at the increased hydrostatic pressures experienced with increasing habitat depth (Pelster, 1997). Third, the lipids may themselves confer buoyancy, which is important as the buoyancy conferred by swimbladder gases decreases with increasing depth (Pelster, 1997). These three explanations offer a strong set of reasons as to why swimbladder lipid mass increases so strikingly with habitat depth in *C. rupestris*. Again, it is unclear at this point how plastic this response to increasing habitat depth is: would any *C. rupestris* individual allocate more lipids to the swimbladder if moved to a deeper habitat with increased hydrostatic pressure, or is this a response only of those individuals who are genetically suited to the bathypelagic?

There is an interesting insight to be had into lipid storage in both the liver and swimbladder when analysing the outputs of the PCA analysis. The Lipid Storage component correlated strongly with liver and swimbladder masses to the slight detriment of most body size variables, and increased significantly with habitat depth. This shows the trends outlined above, where lipid storage in the liver and swimbladder serve to improve survival of food shortages and buoyancy, respectively. The third component of the PCA, though having a relatively small Eigenvalue, was selected because it demonstrated a further interaction between liver and swimbladder lipid storage: that in the bathypelagic, more energy is allocated to the increase in swimbladder lipids than the liver. The implications of this are perhaps that, crucial as lipid storage in the liver is in the bathypelagic, the need for lipid allocation to the swimbladder is greater. This might be because hydrostatic pressure, which mediates all the reasons for increased lipid storage in the swimbladder, increases steadily with increasing depth, where food shortage, the reason behind increased liver energy storage, is consistent throughout the bathypelagic. The selection pressure for increased swimbladder lipid storage therefore becomes more intense as habitat depth continues to increase, where the selection pressure for increased liver lipid storage remains consistent. Though the Energy Allocation component did not have a large magnitude, the trade-off between liver and swimbladder lipid storage is still interesting, and worthy of note when discussing changes in energy storage across habitat depths.

*Conclusion*

Overall, it is clear there is some level of morphotype segregation occurring in *C. rupestris* in response to habitat depth. This segregation can be linked strongly to the meso-bathypelagic transition. There are clear trends for smaller, more elongated bodies, larger eye and mouth sizes, and greater lipid storage in the liver and swimbladder with increasing habitat depth. There is evidence of distinct ‘shallow’ and ‘deep’ morphotypes as demonstrated by the unsupervised Random Forest technique, which showed the creation of

two separate morphological clusters associated with the shallowest and deepest habitat depths respectively (Cluster 3 and Cluster 1). The key selection pressures for the changes observed are mainly associated with the decreased food availability in the bathypelagic and implications for energy efficiency, in addition to the drive for greater buoyancy from the swimbladder at deeper levels, and greater long-distance vision for predation in the bathypelagic.

Though there is the potential for some of the patterns observed to be a response to living at a particular habitat depth as opposed to an adaptation to those depths, namely the increases in liver and swimbladder masses with depth, it is likely that there is some level of genetic predisposition for living at greater depths. These findings are in strong agreement with the conclusions of Gaither et al. (2018), who determined genetic segregation of *C. rupestris* individuals according to habitat depth in line with the meso-bathypelagic boundary. This segregation was linked to the migration of mature individuals rather than the existence of distinct breeding populations. In the future, further support for this theory could be gained by analysing otolith microchemistry. Otolith $\delta^{18}O$ and $\delta^{13}C$ have both been used to demonstrate ontogenic vertical migration in grenadiers previously (Lin et al., 2012), and similar studies could be carried out on *C. rupestris* to confirm when and how individuals migrate to their preferred habitat depths.

For now, with strong evidence of both genotypic and morphotypic segregation by depth in this species, the next step is to uncover the link between genotype and morphotype in *C. rupestris*, and to demonstrate how this relationship results in adaptations that make certain individuals better suited to deeper habitats.

# Chapter 3: Changes to Protein Function Across the Meso-Bathypelagic Boundary in *C. rupestris*

## Introduction

Studying protein structure and function allows us to bridge the gap in our understanding from the genetic to the morphological in *C. rupestris*. Here, I follow on from the work of Gaither et al. (2018), who identified non-synonymous single nucleotide polymorphism (SNP) changes in six functional loci across the meso-bathypelagic boundary in *C. rupestris*. It is important to understand how these genotypic changes might be translated into physiological, morphological, or behavioural adaptations in this species. Here, I use the annotated genome which Gaither et al. assembled to analyse how protein function is affected by the SNPs at those loci, and illuminate a link between genetic segregation, protein function, morphology and behaviour. In order to understand this work, however, it is important to understand the wider context in the literature, and explore how protein function changes with increasing habitat depth in other studies.

*Known protein adaptations to life in the deep sea*

Proteins are sensitive to changes in temperature and pressure. Both these factors change over the habitat range of *C. rupestris*. Hydrostatic pressure steadily increases with depth. Temperature decreases across the *C. rupestris* depth range, though this change is focussed at the top of the range, and levels off quite quickly (Gage and Tyler, 1992; Childress, 1995). Protein function can typically only be sustained within a specific temperature and pressure range, and species adapted to different habitat depths will demonstrate adaptive variations in homologous proteins which allow them to function in the different environmental conditions (Somero, 1998). There are several examples of proteins with well documented differing homologues in species from varying environmental temperature and pressure conditions. It is useful to explore the information available on these proteins, as it allows us

to gain a greater understanding of how protein structure and function can be expected to change across the depth range of *C. rupestris*.

Actin is a protein responsible for contractile movements in the muscle and other cells, and has been well studied in terms of temperature and pressure sensitivity. Reversible polymerisation is an important component of actin function, and this process is affected by changes in both pressure and temperature (Somero, 1998). In a study of actin in deep sea teleost fish, species from the deepest, coldest habitats demonstrated the lowest enthalpy and entropy changes with the polymerisation reaction, meaning less energy was required to initiate the reaction in deeper habitats (Swezey and Somero, 1985). This was linked with a reduced reliance on hydrophobic interactions in the polymerisation, which are associated with large enthalpy requirements and are therefore selected against in the cold deep sea. The same study also explored actin adaptation to pressure. They found that the volume change in the actin polymerisation reaction of the abyssal *Coryphaenoides armatus* was six times lower than in the shallower living congener *Coryphaenoides acrolepis.* Lower volume change in a reaction demonstrates lesser sensitivity to pressure, so the species from the deeper habitat can be seen as pressure-adapted (Somero, 1998). Further study of this protein in *Coryphaenoides* by Morita (2003) showed that pressure adaptation in *C. armatus* can be linked to only three amino acid substitutions, which cause tighter binding to ATP and $Ca^{2+}$ and improve actin structural stability. These findings were supported by Wakai et al. (2014), who further linked the greater stability of ATP domains in pressure-adapted species to the greater number of salt bridges formed.

The dehydrogenases are a family of proteins that have been widely studied in the context of temperature and pressure adaptation in the deep sea. At depths of only 500-1000 m, selection for improved resistance to pressure has been demonstrated (Somero, 1998). This is particularly shallow for hydrostatic pressure to become a constraint; in actin, pressures nearing those at 2000 m are required to necessitate pressure adaptive change (Swezey and

Somero, 1985). In this protein family, as in many enzymes, adaptations to temperature and pressure can create a trade-off. High pressure causes non-adapted proteins to unfold and denature by disrupting intramolecular interactions (Gross and Jaenicke, 1994). Greater rigidity in the protein confers resistance to pressure-induced denaturation, and also increases thermostability. This comes at the cost of catalytic activity, however; cold (but not pressure) adapted proteins have more flexible structures as a result of weakened intramolecular interactions that allow catalysis to occur at a lower energy cost (Gerday et al., 1997). Greater protein flexibility allows improved catalytic activity at low temperatures, but greater rigidity is adaptive for high pressures. The deep sea, with its low temperatures and high hydrostatic pressures, therefore presents confounding selection pressures. It appears that resistance to pressure often wins this trade-off. Lactate dehydrogenase (LDH) in deep sea species demonstrates low sensitivity to pressure along with an increase of activation free energy and enthalpy characteristics, at the cost of catalytic efficiency (Somero and Seibenaller, 1979). This demonstrates the importance of hydrostatic pressure on the function of proteins: adaptation to high pressure is necessary to preserve function in the deep sea, even at the cost of activity. Further exploration of LDH demonstrates that adaptation to deep habitats can take place with a relatively low number of amino acid changes. Brindley et al. (2008) demonstrated that when comparing LDH from the deep *C. armatus* and the shallow Atlantic cod, *Gadus morhua*, only 21 amino acid substitutions took place, which nonetheless permitted significantly higher thermostability and resistance to high pressures in *C. armatus*. The authors highlighted replacements of lysine residues, particularly changing to arginine, as important in conferring stability to the *C. armatus* proteins by increasing structural rigidity, again at the cost of catalytic activity.

While actin and dehydrogenases are probably the most thoroughly studied proteins in terms of depth adaptation, the impacts of hydrostatic pressure and temperature are the same for all body proteins, and new study proteins are still coming to the fore. For example, Porter et al.

(2016) recently demonstrated that the light-sensitive G-protein coupled receptor (GPCR) protein opsin was less compressible in fish and cephalopod species from deeper habitats compared with those from shallow habitats. This adaptation was linked to a small set of amino acid sites which underwent stabilising selection in deep sea species in response to increased hydrostatic pressure. There is undoubtedly a lot more to discover about intrinsic protein adaptations to increasing habitat depths, but there are a few common concepts which are repeated throughout the existing studies. First, hydrostatic pressure and temperature are key selection pressures at play; second, adaptation to pressure results in increased protein rigidity, sometimes to the detriment of protein activity; and third, relatively few amino acid changes are required to confer protein adaptation to increased habitat depth.

*Initial exploration of potential study proteins*

In order to explore how habitat depth might affect protein structure and function in *C. rupestris*, I used the data from Gaither et al. (2018), who recently reported six functional gene loci which encounter fixed SNPs in fish from deeper than 1500 m in this species. Modelling the proteins these genes code for would allow me to examine the effects the substitutions had on function, and therefore discuss the adaptive benefits, and potentially link those benefits to the morphological trends I have already reported. The proteins in question have a range of functions (listed in Table 9), and could therefore provide a real insight into a variety of adaptive processes that occur with depth.

The protein sequences for the six proteins affected by SNPs were obtained, and I ran these through ModWeb, a web server which calculates structural models for protein sequences based on template structures recorded in the Protein Data Bank (PDB; Berman et al., 2002). This allowed me to ascertain which of the proteins had structures similar to other known protein structures. Greater similarity to known protein structures means that the homology models calculated by ModWeb are more accurate, so more complete analyses can be carried out using the models.

Table 9: Six proteins identified as having amino acid substitutions in *C. rupestris* individuals from deeper habitats by Gaither et al. (2018), and their functions.

| Protein name | Function |
|---|---|
| **ROCK1** | Rho-associated protein kinase associated with actin cytoskeleton regulation and maintenance of cell polarity. Used in smooth muscle contraction, cell adhesion and motility, stress fibre formation and more (UniProt entry Q13464; UniProt Consortium, 2018). |
| **EGFR1** | Epidermal growth factor receptor essential in development, regulator of cell proliferation, survival and migration. Overexpression is linked to epithelial tumour growth (Fromm et al., 2008). |
| **YOF5** | Oxidoreductase activity (UniProt entries C0HB08 and D3PHE9; UniProt Consortium, 2018). |
| **ATG9** | Trans-membrane protein involved in generation of autophagosomes as part of cell autophagy (Ungermann and Reggiori, 2018). |
| **OBSL1** | Obscurin-like 1, involved in structural regulation of the M-band in muscles (Geisler et al., 2007). |
| **CAC1E** | Voltage-sensitive calcium channel regulating entry of ions into calcium-excitable cells. Involved in muscle contraction, gene expression, hormone/neurotransmitter release, and cell motility, division, and death. Alternative name CACNA1E (UniProt entry Q15878; UniProt Consortium, 2018) |

EGFR1 retrieved homology matches with good sequence identities. These included the human EGFR protein (55% identity, PDB entry 4UV7; Lim et al., 2016), which matched to the first half of the *C. rupestris* EGFR1 protein, and a human ErbB4 kinase domain (PDB entry 2R4B; Wood et al., 2008) which had a high sequence identity match at 79%, but only covered a small amount of the second half of the protein. Overall, though high sequence identities were found in several homologous structures, it was decided that this protein was too large, with too many different and complex domains, to effectively analyse in the context of this project.

YOF5 yielded some homology models with fairly strong sequence identity matches, including a eukaryotic elongation factor (65% identity, PDB entry 2N51; Wu et al., 2016), and a viral hydrolase (34% identity, PDB entry 4cpy; Escuret et al., 2014). These matches had very poor coverage of the *C. rupestris* YOF5 sequence, however, meaning a homology model of YOF5 could only be generated for small segments of the protein. Furthermore, the functions of the protein matches in this case were so different from the function of YOF5 that it was clear that the models generated could not be used for further analysis. For ATG9,

the outlook was even worse, as no reliable models could be generated by the ModWeb server. CAC1E demonstrated similarly poor homology model matches. The highest sequence identity match was 32% (PDB entry 5JR0; Niitsu et al., 2018), but the coverage for this match was so poor that only a single alpha helix was modelled. Most of the protein had no matches whatsoever, with only two additional matches with very poor coverage and identity. CAC1E was therefore also excluded from analysis.

ROCK1 was the protein that, in Gaither et al.'s (2018) study, looked the most promising for demonstrating interesting adaptations to habitat depth. This was because three distinct non-synonymous SNPs were found close together in the gene sequence: a finding that strongly implied the SNPs were the result of active selection, rather than drift. Unfortunately, the homology models generated did not suggest that ROCK1 should be explored further using the methods I had planned. Though there was a very strong 85% identity match in the first section of the protein sequence to a bovine ROCK protein (PDB sequence 2F2U; Yamaguchi et al., 2006), the rest of the protein had very poor matches. The match with the greatest coverage had only 10% sequence identity with ROCK1, and was from a dynein motor domain (PDB sequence 3VKG; Kon et al., 2012). That such poor structural matches were found to generate the homology model suggested that it would not be accurate, and that further analysis of the ROCK1 model in this context would be futile.

Finally, OBSL1 yielded the best homology matches. A homology model was generated from a section of the giant protein titin (PDB entry 5JDD; Bogomolovas et al., 2016). The sequence identity was satisfactory at 31%, and the match had nearly 100% coverage of the inputted protein sequence. Furthermore, the type of structure that OBSL1 takes, with linked domains composed of only beta sheets, is simpler to model than more complex structures such as EGFR1. For this reason, OBSL1 was taken as the focal protein for further analysis.

*OBSL1 and its function*

OBSL1, or obscurin-like 1, is a large filament protein with a structural role in the muscle. In skeletal muscle, muscle fibres are long, multinucleated cells which contain thin, hair-like structures called myofibrils (Squire et al., 2005). These myofibrils are arranged parallel to each other, and the bundles of myofibrils can be broken up into contractile units called sarcomeres, which are visible with an electron microscope (Figure 13). The sarcomere is composed of thick filaments containing the protein myosin, and thin filaments containing actin. The thin filaments are anchored at each end of the sarcomere in a region called the Z-disk, which is visible as a dark stripe in Figure 14. The thin filaments overlap with the thick filaments in the middle of the sarcomere. This region of overlap is called the A-band. At the centre of the A-band is the M-band (Figure 14), which is where myosin rods are arranged in a characteristic anti-parallel manner. This band appears dark because it contains many cross-linking proteins which maintain myosin structure (Sauer, 2012). OBSL1 is one such protein.

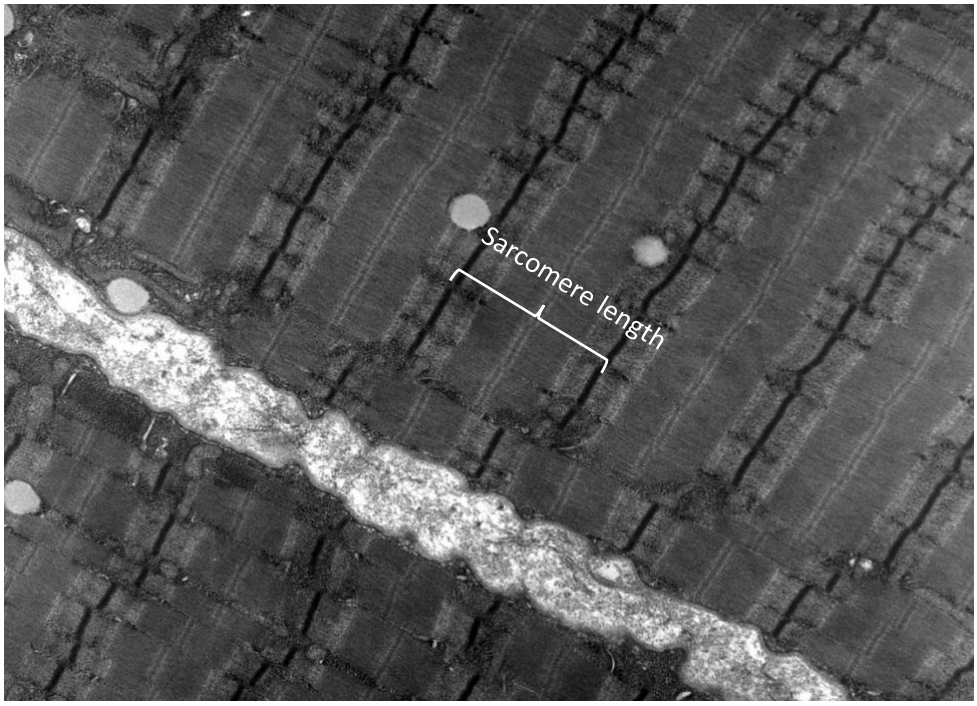


Figure 13: Transmission electron microscope image of human skeletal muscle. The visible striation marks repeating sarcomere units. The length of one sarcomere unit is labelled. Image was modified from https://commons.wikimedia.org

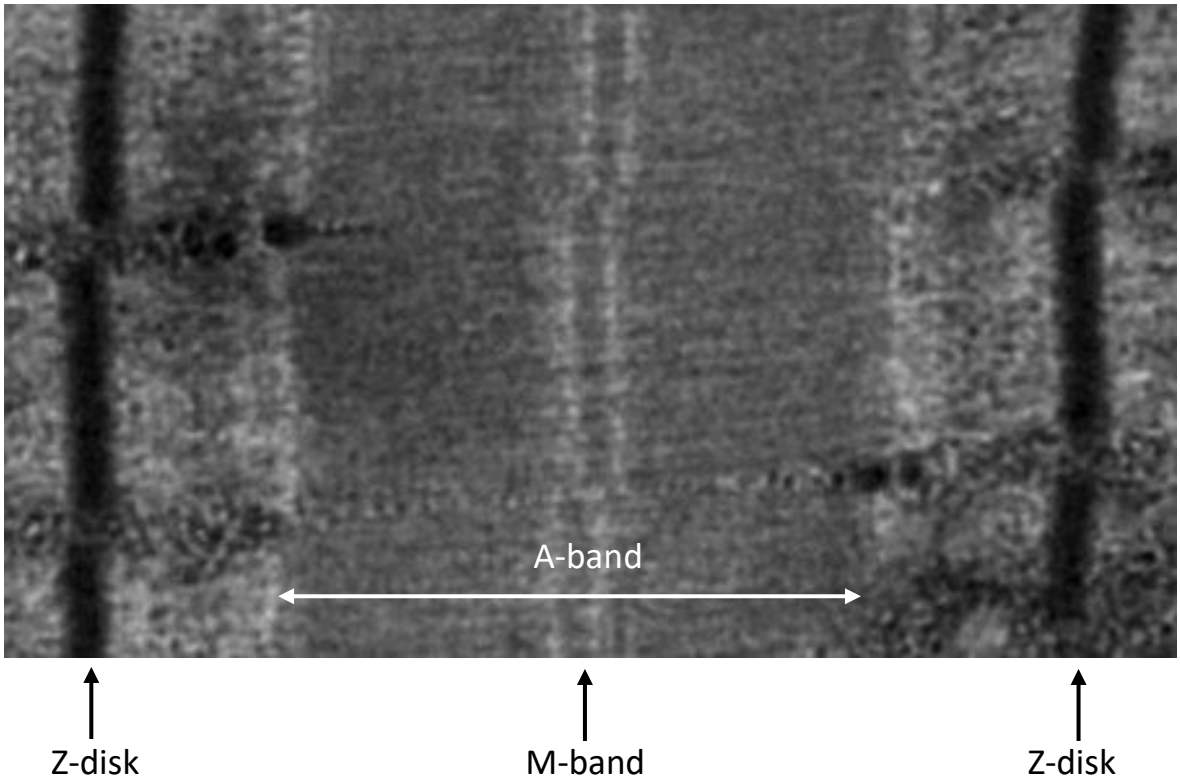


Figure 14: Magnified image of the sarcomere highlighted in Figure 13. The Z-disk, A- and M-bands are marked.

Myomesin is the main protein believed to be responsible for maintaining structural links between the thick filaments in the M-band, holding them in a hexagonal lattice formation. OBSL1 and its close homologue obscurin both interact with myomesin, as well as the C-terminal M10 domain of titin, a giant elastic protein which connects the Z and M-bands (Lange et al., 2019.) As shown in Figure 15, the positions of these proteins relative to one another and their binding sites have been graphically mapped out in two dimensions, however no holistic, 3D model has been created to demonstrate how these proteins fit into the M-band and interact with the thick filaments to maintain muscle structure. For this reason, the impact of potential changes to OBSL structure and function must be assessed by extrapolating the known information about the function of the M-band itself, as well as the role of the other M-band proteins. Lange et al. (2019) demonstrated that the M-band is important not only for maintaining structure, but for managing force imbalances that occur during muscle contraction, essentially acting as a shock absorber.

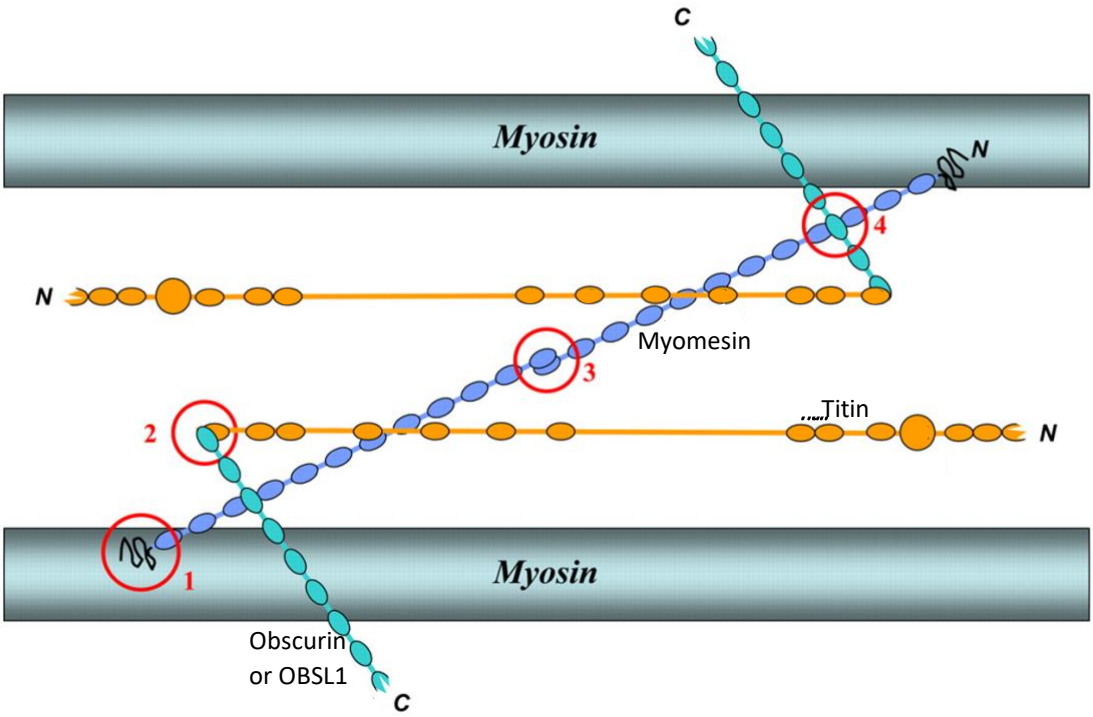


Figure 15: Representation of the interactions between structural elements of the M-band, modified from Fukuzawa et al. (2008) Fig. 8. OBSL1 (or obscurin), myomesin and titin are shown in cyan, purple and orange respectively. Protein domains are represented by ovals, and the N- or C-terminals of the proteins marked. Known protein interactions are marked with red circles: (1) myomesin domain My1 and myosin; (2) OBSL1/obscurin Ig1 with titin M10 domain; (3) myomesin antiparallel heterodimer on domain My13; (4) OBSL1/obscurin Ig3 with myomesin My4-My5.

OBSL1 is composed of repeating immunoglobulin (Ig) domains, which are themselves made up of a "sandwich" of beta sheets (see Figure 16). Titin and myomesin have similar structures, though both have additional domain types, namely fibronectin domains (Agarkova and Perriard, 2005; Pernigo et al., 2010). The Ig domains serve as binding sites, but also afford the protein a certain level of elasticity. Studies of titin and myomesin Ig domains have shown they are able to unravel slightly before springing back to re-form (Schoenauer et al., 2005). Similar qualities can be expected in OBSL1, which fulfils an analogous role in the M-band, and which is made up of similar Ig domains. This elasticity is no doubt useful in the shock-absorbing capacity in which the M-band functions (Lange et al., 2019).

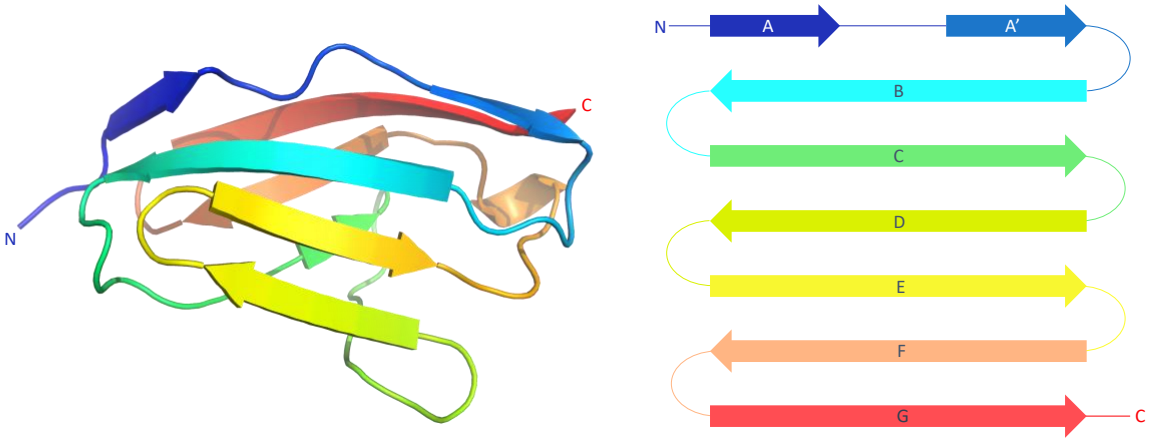


Figure 16: Model of a typical OBSL1 Ig domain, with a colour coded representation of the beta strand labels. The N- and C-terminal ends of the domain are labelled. Note the characteristic 'sandwich' of beta sheets.

OBSL1, as a structural protein, might face significant changes depending on the impact of the SNP discovered by Gaither et al. (2018). These changes might in turn have an impact on morphology or behaviour, as OBSL1 is involved in muscle function. It was hypothesised that OBSL1 would follow the trends observed in actin and the dehydrogenases, with SNPs causing an increase in rigidity as a response to increased hydrostatic pressure in deeper habitats. In line with this was the expectation that increased rigidity might impact the shock-absorbing function of the M-band. Exploration of these two hypotheses would allow us to draw conclusions about the mechanisms behind adaptations to different habitat depths in *C. rupestris*.

## Methods

### *Initial analysis*

The first step in exploring the role of the known SNP in the function of OBSL1 was to analyse the homology model generated by ModWeb. The coordinate file was downloaded and the structure analysed in two programs: PyMol (DeLano, 2002), and Coot (Emsley et al., 2010). PyMol was used to analyse gross structural forms and the location of the

substituted amino acids, whereas Coot allowed for a more finescale analysis, and for the changes caused by the amino acid substitution to be explored in more detail.

It was important to establish whether the amino acid substitution would have any impact on binding. Though there are no known binding partners for the specific Ig domain on which the substitution is found, the incomplete knowledge of the structural layout of the M-band (Lange et al., 2019) means that the existence of unknown interactions cannot be ruled out. The PDBe website was used to search for any OBSL1, obscurin, titin, or myomesin Ig domains with recorded structures demonstrating how they bind to other proteins. These structures were then downloaded, and the Ig domain affected by the SNP was aligned to the structure of the binding model Ig domain using PyMol. This was done to see whether the amino acid substitution might affect the protein binding surface.

*OBSL1 size analysis*

The OBSL1 homology model consisted of two and a half Ig domains, and had full coverage of the inputted protein sequence. Since the literature shows that OBSL1 is a long protein, and an incomplete Ig domain was formed, it was clear that the initial protein sequence used was incomplete. Establishing the length of the whole protein was necessary to explore additional non-synonymous SNPs in the genome not featured in Gaither et al. (2018), and to analyse whether these had any further impacts on function that might be adaptive.

First, the Ig domain sequence affected by the SNP was run through UniProt BLAST (UniProt Consortium, 2018), which identified protein sequences with high sequence identity to the search sequence. When the domain was matched to a protein, it was matched to multiple places on that protein, representing multiple Ig domains. These multiple domains could then be used to search for the related gene sequences in the *C. rupestris* genome, however the protein sequences first needed to be converted into gene sequences. For each domain match, the protein sequence of the match was taken and run through NCBI protein BLAST (Madden, 2002). This allowed identification of the protein and gene sequences for the whole

protein from which the matched domain originated. The whole gene sequence was taken, and run through the GeneWise protein-gene pair alignment tool (Madeira et al., 2019), along with the protein sequence of the matched domain. This allowed the matching of the relevant gene sequence for that particular domain. This gene sequence was then run through the NCBI BLAST website against the *C. rupestris* genome that was uploaded there by Gaither et al. (2018). When carried out systematically, this process allowed the discovery of the gene sequences of many different OBSL1 Ig domains. These were mapped onto the contig on which the OBSL1 gene was found to demonstrate the total size of the protein.

After the full OBSL1 protein sequence was identified, it was necessary to find all coding regions for this protein, not just regions that coded for the Ig domains. Though some coding regions had been identified on the contig in question by Gaither et al. (2018), and were accessible through the genome browser JBrowse (Buels et al., 2016), some of the identified coding regions for Ig domains were outside the exons identified on that platform. It was therefore important to establish where all of the exons for this protein were. The gene sequence for the entire section on the contig in which gene sequences for Ig domains had been found was downloaded and run through the exon prediction program Fgenesh (Solovyev et al., 2006). This program predicted the coding regions for OBSL1 within the genome.

*SNP modelling*

With the full size of the protein established and the coding regions in the genome found, it was then possible to ask the question of whether any additional non-synonymous SNPs were present across the meso-bathypelagic boundary in this species. A file was made with SNP identities for all of the 60 genomes constructed by Gaither et al. (2018) for the whole contig on which OBSL1 was found. These took the form of ambiguity codes for each copy of the nucleotide affected by a SNP for each individual. These ambiguity codes were then colour coded, and the genomes sorted by habitat depth. This allowed the visualisation of whether

any SNP became fixed at the 1500m boundary identified by Gaither et al. (2018). Any fixed nucleotide changes within OBSL1 exons were transferred to a new data file for further analysis.

Each SNP then had to be analysed for whether it was synonymous, resulting in the same amino acid being incorporated into the protein at all habitat depths, or non-synonymous, resulting in an amino acid substitution. For each fixed SNP, the gene sequence for the domain in which it was found was run through the translation program ExPASY (Gasteiger et al., 2003), which output the protein sequence. The same process was carried out a second time, but the nucleotide affected by the SNP was changed to the deep habitat form. A pair alignment of the two protein outputs could then be run using the EMBOSS Water pairwise alignment tool (Madeira et al., 2019). If the SNP was synonymous, the two protein sequences would have 100% identity, but if the SNP was non-synonymous, the affected amino acid could be identified and recorded. This process was carried out for all fixed SNPs within OBSL1 coding regions of the contig.

Once the non-synonymous SNPs were identified, it was necessary to analyse the structural and functional impact of the resultant amino acid substitutions. Each domain affected by a non-synonymous SNP was run through ModWeb to generate a homology model. These models were then analysed using PyMol and Coot, as described above. This allowed a complete investigation of the impact of all non-synonymous SNPs that occur across the 1500m depth boundary in *C. rupestris*.

In order to compare functional changes in other related proteins with depth, the gene sequence and SNP list was also obtained for the protein myomesin, which performs a similar role in the M-band. The SNP list was colour coded and analysed for fixed and non-synonymous SNPs using the same method as described for OBSL1.

## Results

*Initial analysis*

Exploration of the initial OBSL1 protein sequence showed that, on the 2.5 linked Ig domains of the model, the SNP caused an amino acid substitution at an arginine residue, ARG 172, which is found on the middle domain. This arginine residue, which is altered to a lysine in the deep living fish, is located just after the C-terminal end of beta strand G, facing outwards from the domain (Figure 17). The substitution of arginine to lysine cannot be expected to result in a large functional or structural change in this protein. Both amino acids have long, flexible, positively charged sidechains capable of forming ionic bonds, H-bonds, and van der Waals interactions. The only functional difference is that the arginine sidechain is slightly more complex, and has the potential to form a greater number of interactions. This might be relevant if the substitution were to take place on a surface of the domain used for binding to other proteins or structures. Known interactions between similar domains were therefore analysed to explore whether this substitution takes place on a surface used for binding in homologous domains.

Figure 17: Homology model of the initial OBSL1 sequence, displaying the 2.5 Ig domains that were generated. The SNP identified by Gaither et al. (2018) causes an amino acid substitution in ARG 172, which is highlighted in red and whose sidechain is represented by a wire structure, facing outwards from the domain. This arginine residue is replaced by a lysine, whose sidechain is similarly long and positively charged.

Two main homologous protein domains demonstrated interesting results when this OBSL1 domain was aligned to them and binding was analysed. The first domains were from the N-terminal of titin, a doublet named Z1Z2. Two Z1Z2 doublets will form an antiparallel

sandwich complex with telethonin, a ligand protein found in the Z-disk of the sarcomere ( PDBe entry 1ya5; Zou et al., 2006). The OBSL1 domain affected by the SNP aligned onto the Z1Z2 domains so that the affected amino acid faced inwards, into the binding interface, on two out of four titin Ig domains in the sandwich complex (Figure 18). Second, the model demonstrating the interaction between titin and OBSL1 itself was analysed (PDBe entry 2wp3; Pernigo et al., 2010). The titin domain M10 interacts with the OBSL1 domain OL1, forming a V-shaped complex. When the OBSL1 domain affected by the SNP was aligned to domain OL1, the affected amino acid is on the same plane as the binding surface, but at the 'open' end of the V-shaped complex (Figure 19). No further analyses could be conducted, as the SNP does not affect the actual binding surface, but it was useful to establish that the amino acid change took place on a surface that could be used for binding interactions.

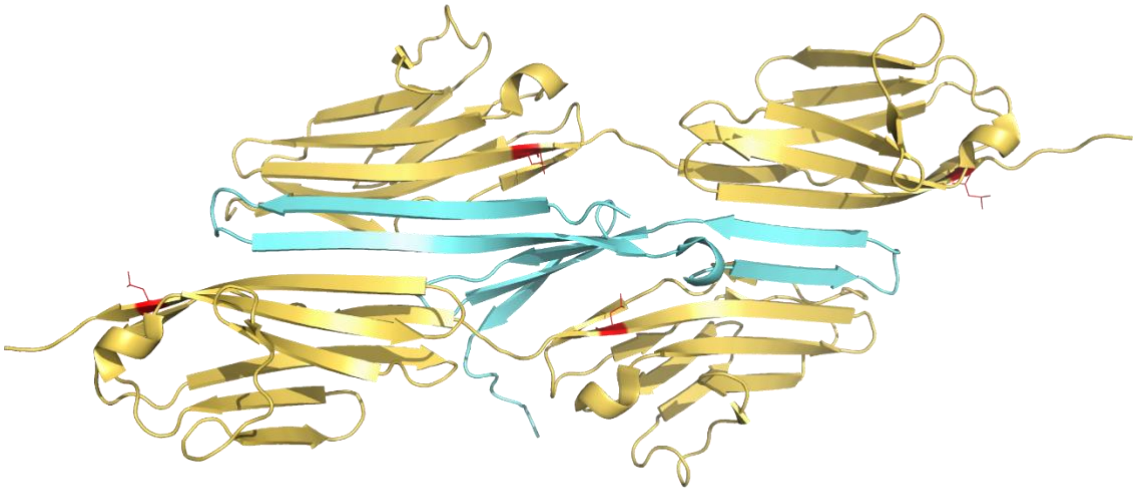

Figure 18: Model of the sandwich complex formed between two Z1Z2 titin domains (yellow) and telethonin (blue) in the sarcomeric Z-disk. When the OBSL1 domain affected by the SNP identified by Gaither et al. (2018) was aligned to each of the titin Ig domains, the affected amino acids were aligned to the relative titin amino acid. Those residues were highlighted in red on the titin domains and their sidechains represented by wire structures, demonstrating that, in two of the domains, an amino acid substitution there might affect binding to telethonin.

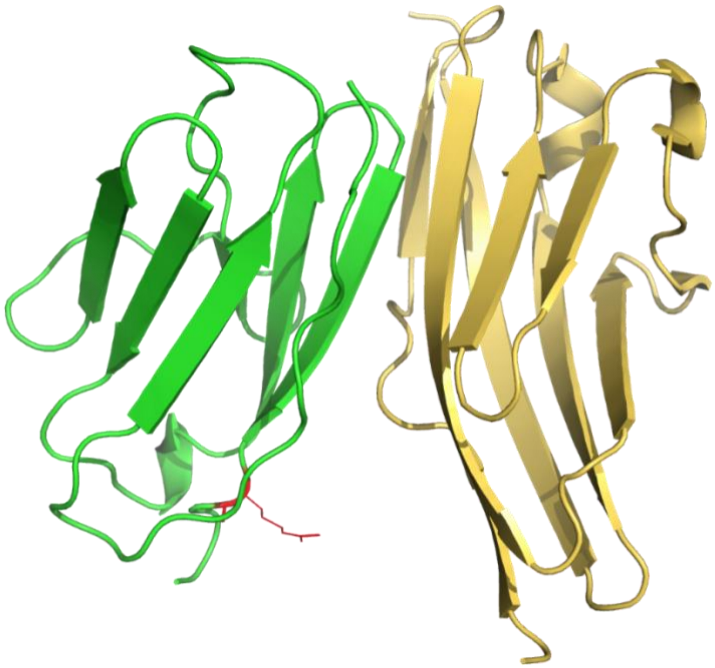

Figure 19: Representation of how the SNP identified by Gaither et al. (2018) might affect OBSL1 (green) interaction with the titin M10 domain (yellow). The model for the V-shaped complex between OBSL1 Ig1 and titin M10 was found, and the OBSL1 domain affected by the SNP was aligned onto the OBSL1 Ig1 domain from the model. The amino acid affected by the SNP is highlighted in red, the sidechain represented by a wire structure.

*OBSL1 size analysis*

Matches from UniProt BLAST were found in an uncharacterised protein from rainbow trout, *Oncorhynchus mykiss*, and an obscurin-like isoform from Atlantic salmon, *Salmo salar*. Identity matches were 84.0% and 83.3%, respectively. When the multiple Ig domains from these proteins were aligned to the *C. rupestris* genome using NCBI BLAST, a total of 26 potential additional domains were found on the same contig on which the domain affected by the SNP was found. Most of these domains were identified with matches from homologous domains from both *O. mykiss* and *S. salar*. Overall, the genome showed coding regions for Ig domains across a 36,000 bp region, and when the predicted exons were retrieved from Fgenesh, all predicted Ig domains were found within predicted coding regions. A map of the contig in question, with markers for predicted Ig domains, can be seen in Figure 20.

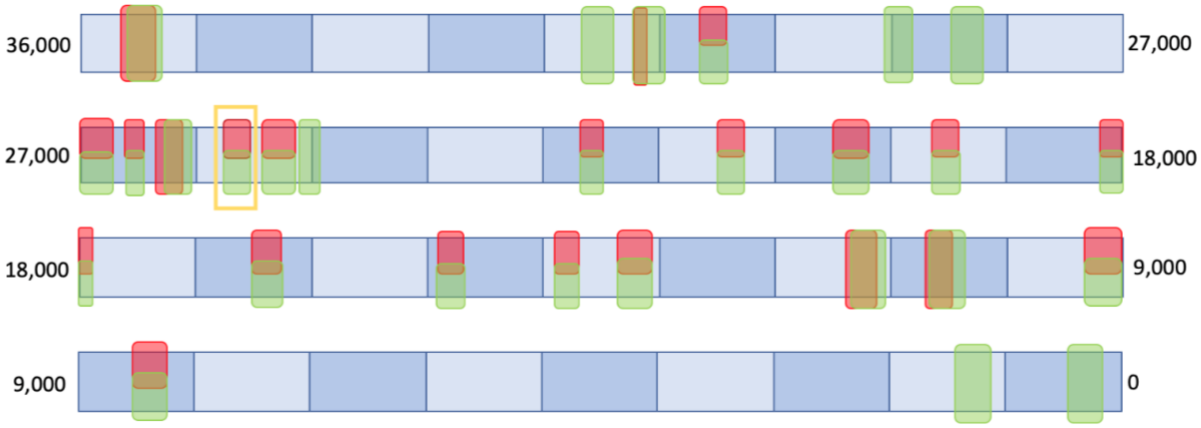


Figure 20: Map of coding regions for OBSL1 Ig domains on the *C. rupestris* contig as found from homology matches from *S. salar* (green boxes) and *O. mykiss* (red boxes). The domain on which the original SNP was located by Gaither et al. (2018) is highlighted in the yellow box.

*SNP modelling*

Upon obtaining the total list of SNPs in the OBSL1 contig from the 60 genomes sequenced by Gaither et al. (2018), the list was structured according to fish habitat depth and colour coded to identify any nucleotide changes which were fixed at a certain depth. Across the 36,000 bp sequence on which OBSL1 coding regions are found, a total of 672 SNPs were detected. Of these, 245 were determined to be fixed. This fixation took place in most cases between the fish from 1500 m depth, and those from 1800 m. The three shallowest depth groups, 750, 1000, and 1500 m, generally showed a mix of nucleotide identities, both homo- and heterozygous, where at 1800m, identities became fixed to a homozygous form, often different to the homozygous forms found in shallower habitats. When the set of fixed SNPs was narrowed to only those found within identified Ig domains and exons, 72 were found to be fixed and within coding regions. Finally, when those 72 exons were analysed for being synonymous or non-synonymous, a total of 32 non-synonymous, fixed SNPs were found within coding regions of OBSL1. A table illustrating the position of these SNPs across the contig and the nature of their fixation can be found in Appendix 16.

With the 32 non-synonymous SNPs identified, their domains were run through ModWeb to generate homology models for structural analysis. Because ModWeb requires the existence of homologous structures in its databanks to create structural models, not all the non-

synonymous SNPs could be analysed, as when their domains were fed into PyMol, the homology model generated did not cover the portion of the inputted sequence containing the changed amino acid. Because of this, only 25 of the non-synonymous SNPs could be analysed in reference to a modelled structure. See Table 10 for a full list of effects of the SNPs on the function of OBSL1.

When the same analysis was carried out on myomesin, no SNPs were found to be fixed at any depth level whatsoever. This meant that neither further analysis of synonymous or non-synonymous identity, nor structural modelling, could be carried out.

Table 10: List of amino acid changes and their effects from the non-synonymous SNPs found to be fixed at a certain depth across the coding regions of the OBSL1 gene. 'Shallow' and 'deep' amino acid headings refer to the identity of amino acids affected by the SNPs at 750-1500m, and 1800m, respectively. SNPs are numbered according to the order in which they affect the protein, with 1 at the N-terminal end, and 26 at the C-terminal end. Table continues on pages 72 and 73.

| SNP | Shallow amino acid | Deep amino acid | Analysis |
|---|---|---|---|
| **1** | Leucine | Methionine | Amino acid change takes place just before the N-terminal of the generated homology model. This is the first Ig domain in OBSL1, and the change from the simple aliphatic leucine to methionine, able to form strong disulphide bonds, may be significant, though its effect could not be ascertained. |
| **2** | Alanine | Valine | Both amino acids face outwards from the domain. Neither is charged or polar. The shape of the protein surface would be changed however, since alanine has a much smaller sidechain. If involved in binding, the binding surface may be affected. |
| **3** | Alanine | Serine | Affected amino acid faces outwards from the Ig domain and is found near the C-terminal end of the G strand. Change to serine allows H-bonds to form where they cannot in alanine, however there is no known binding partner for this domain, and no interactions within the protein seem to be made possible by the addition of the OH group. |
| **4** | Lysine | Methionine | Amino acid sidechain faces outwards from the Ig domain. Change from lysine to methionine removes potential for both ionic and H-bonds with other proteins in deeper fish. Methionine is also more hydrophobic than lysine. As the sidechain is modelled facing outwards, the addition of methionine may alter this structure slightly as the hydrophobic sidechain will be more attracted to the core of the protein. |
| **5** | Asparagine | Lysine | Asparagine on the A strand can form an H-bond with 75 GLN from the adjacent G strand. The change to lysine makes this impossible as its sidechain is too long, and makes the bond distances too great. This suggests lesser forces are holding the Ig domain together in the deeper form. Although new inter-protein interactions might be made possible by the change to lysine, no binding partners are known for this domain. Same domain as SNPs 6 and 7 (Figure 23). |
| **6** | Cysteine | Arginine | The cysteine residue is on the E strand and faces inwards on the Ig sandwich, and may form a disulphide bond with either the B strand 13 CYS, or the F strand 63 CYS. This would bind the two opposing sides of the Ig sandwich very tightly together. With the change to arginine, its long sidechain is also modelled as extending into the domain, potentially interrupting the internal structure of the Ig domain. Same domain as SNPs 5 and 7 (Figure 23). |
| **7** | Glutamic acid | Alanine | Glutamic acid on the G strand can dorm an H-bond with the OH group from 4 THR on the A' strand. This may hold the beta sheets together more securely in the shallower living fish in a way that is not possible when the glutamic acid is replaced with alanine in the deeper living fish. Same domain as SNPs 5 and 6 (Figure 23). |
| **8** | Threonine | Alanine | Though threonine has an OH group which alanine lacks, no functional change was discernible with this substitution, as both amino acid sidechains face outwards from the domain, and do not appear to interact with anything. Same domain as SNP 9. |
| **9** | Arginine | Serine | The arginine sidechain faces inwards on the domain, and is located on a linker region between beta strands. It is able to form an H-bond with the OH group from 50 TYR, linking the C and D chains more tightly. The change to serine removes this possibility, and as it is much smaller, might create a gap in the internal structure of the Ig domain. Same domain as SNP 8. |
| **10** | Threonine | Alanine | Though threonine has an OH group which alanine lacks, no functional change was discernible with this substitution. |

Table 10 continued.

| | | | |
|---|---|---|---|
| **11** | Valine | Isoleucine | Both amino acids are simple aliphatic residues, and this substitution does not appear to have any functional impact. Same domain as SNPs 12 and 13 (Figure 21; 22). |
| **12** | Methionine | Arginine | Methionine is modelled as facing outwards from the domain, and doesn't appear to interact with anything. It may be part of a binding surface with an unknown protein. When arginine is substituted, a strong positive charge is added as well as the potential for H-bonding. If this is an interaction surface, this would alter the function strongly. Same domain as SNPs 11 and 13 (Figure 21). |
| **13** | Lysine | Methionine | The lysine residue is located next to another, creating an area of positive charge on the protein surface which is interrupted when changed to methionine. The lysine is also able to form an H-bond with nearby 25 GLU, potentially creating stronger intra-domain bonding in the shallower form of the protein. Same domain as SNPs 11 and 12 (Figure 21). |
| **14** | Arginine | Threonine | Arginine faces outwards from the domain, and has a long, positive sidechain. When this is replaced by the much shorter and polar threonine sidechain, the domain surface changes significantly if used for binding. Threonine may form an H-bond with nearby 29 ASP. Same domain as SNP 15. |
| **15** | Arginine | Lysine | The same amino acid substitution, in the same place in the domain protein sequence, as SNP 17. The arginine on the G strand may form an H-bond with both the COOH group from the A strand 11 ASP, or that from the A strand 13 GLU. Either of these bonds might be possible after the change to lysine in the deeper living fish, however not both. This may weaken the intra-domain binding in the fish from deeper habitats. Same domain as SNP 14. |
| **16** | Isoleucine | Threonine | When threonine is substituted for isoleucine on the A strand, H-bonds become possible between the OH groups of either 27 ASN or 25 GLU, both on the B-strand. These interactions are not possible with isoleucine. Same domain as SNP 17. |
| **17** | Arginine | Lysine | The same amino acid substitution, in the same place in the domain protein sequence, as SNP 15. The arginine on the G strand may form an H-bond with both the COOH group from the A strand 11 ASP, or that from the A strand 13 GLU. Either of these bonds might be possible after the change to lysine in the deeper living fish, however not both. This may weaken the intra-domain binding in the fish from deeper habitats. Same domain as SNP 16. |
| **18** | Threonine | Valine | Threonine has an OH group facing outwards, with the potential for external interactions, but is not in a position to form H-bonds with any residues within the Ig domain. Valine is completely aliphatic, so removes the potential for external H-bonds. |
| **19** | Alanine | Proline | The SNP is on a loop between beta strands. The change from alanine to proline would make the domain structure more rigid, and less flexible. |
| **20** | Tryptophan | Arginine | Tryptophan, on the E chain, has a double aromatic ring sidechain, with some potential to form an H-bond with the O from nearby 49 ASN on the B chain. There are several other aromatic residues nearby: 87 HIS, 84 HIS, and 76 HIS. When changed to arginine, distances become too great for H-bonding. Both tryptophan and arginine face outwards from the domain, but have very different properties, potentially causing a large change in binding surface. Same domain as SNP 21. |
| **21** | Threonine | Alanine | Threonine, on the G strand, can form an H-bond with the backbone N from nearby 31 PHE on the A strand. Alanine cannot form such a bond. The domain may be held together tighter in the shallow form. Same domain as SNP 20. |

Table 10 continued.

| | | | |
|---|---|---|---|
| **22** | Methionine | Isoleucine | No significant interactions were observed for either amino acid. Both are modelled as facing outwards from the domain. Same domain as SNP 23. |
| **23** | Glutamine | Histidine | Glutamine, on the E chain, may form an H-bond with nearby D chain 49 LYS. Histidine cannot form such a bond. Both residues are modelled as facing outwards from the Ig domain. Same domain as SNP 22. |
| **24** | Methionine | Leucine | Methionine faces inwards on the domain and is surrounded by multiple aromatic residues: 48 HIS, 72 TYR, 37 HIS, and 35 TRP. Interactions between the S from methionine and the aromatic rings may add structural stability to the domain. This stability may be lost when the change to leucine is made in the deeper living fish. |
| **25** | Cysteine | Arginine | Cysteine faces sideways on the domain, and does not appear to interact with any nearby residues. The change to arginine is significant: where cysteine is polar, arginine has a positive charge. Arginine does not appear to interact with any other residues from the domain itself either, so any functional change would have to concern external interactions and binding. Same domain as SNP 26. |
| **26** | Alanine | Threonine | Change to threonine allows the potential for H-bonds with 53 ARG or the backbone N of 54 GLU. These interactions are not possible with alanine in the shallower form. |

## Discussion

I have identified *C. rupestris* OBSL1 as a long, multi-Ig-domain protein in line with descriptions from the literature. This protein shows a consistent pattern for fixed amino acid changes with increasing habitat depth. The significance of this finding is shown even more clearly when compared to the similar analysis of myomesin, which showed that no identifiable SNPs became fixed with increasing habitat depth at all. That in comparison, 72 fixed SNPs were found in OBSL1, 32 of which resulted in non-synonymous amino acid changes, suggests that the changes occurring in this protein are somehow adaptive. Here, I explore what the nature of this adaptation may be, and how it might align with other examples of protein depth adaptation as well as the morphological findings I presented in Chapter 2.

First, it should be noted that of the two known binding sites of OBSL1, no SNP caused a known change in binding. Though SNP 1 affected Ig1, the domain which binds to the C-terminal titin domain M10 (Pernigo et al., 2010), the changed amino acid was not located on the Ig domain proper, but on the N-terminal region before the Ig domain begins, and would have no discernible impact on OBSL1-titin interactions. The other known interaction of OBSL1 is with myomesin, at Ig3. No SNPs were found on this domain at all. This does not necessarily mean that no binding is affected by the identified SNPs, however. The initial analysis of OBSL1 and alignment with known binding structures of other Ig domains demonstrated that SNPs affect domains on surfaces that could be used for binding (Figure 18; Figure 19). Further, many of the changes analysed in the SNP modelling affected amino acids which faced outwards from the Ig domain, with some causing extreme changes to the shape and chemical properties of the domain surface (e.g. SNP 12; Figure 21). The M-band, where OBSL1 is localised, has a complex structure which has not been comprehensively modelled (Lange et al., 2019), so it is possible that more protein interactions exist with OBSL1 which have not yet been identified. Beyond structural M-band proteins, it is also

possible that OBSL1 interacts with other cellular components. It is therefore impossible to conclude that no SNPs affect extra-protein interactions in OBSL1, because relatively little is still known about its potential binding sites. Although there are many Ig domains with SNPs which cause changes on potential binding surfaces, no conclusive analysis can be carried out without improved understanding of how OBSL1 interacts with other proteins and components in the M-band.

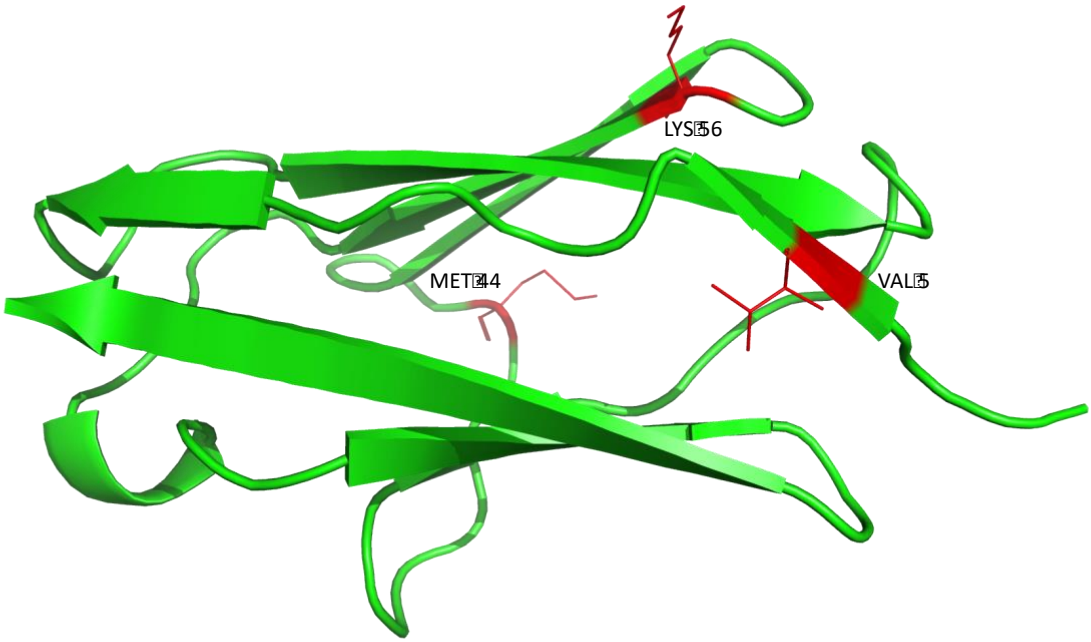


Figure 21: OBSL1 Ig domain affected by SNPs 11, 12 and 13. SNP 11 causes valine (VAL 5) to change to isoleucine; SNP 12 causes methionine (MET 44) to change to arginine; and SNP 13 causes lysine (LYS 56) to be changed to methionine. Affected amino acids are highlighted in red and their sidechains shown with wire structures.

Some amino acid changes caused by the SNPs cause changes which do not appear to influence either protein structure or function very strongly. For example, in SNP 11, one simple aliphatic residue, valine, is substituted for another, isoleucine, with no discernible impact on function (Figure 22). As aliphatic residues contain only carbon and hydrogen in their sidechains, the binding potential of the amino acid does not change from the shallow to the deep form. The small difference in sidechain shape does not appear to have an effect, as the residue faces outwards from the domain and is not close enough to any other structures to interact. No functional change could be detected in SNPs 1-3, 5, 8, 10-12, 14, 18, 20 and 22 either, although the potential for an effect cannot be fully ruled out for several of these, as the dynamics of external interactions and binding partners are not known for all domains.

This finding is intriguing, as the very high number of non-synonymous SNPs in OBSL1 compared to the functionally similar protein myomesin suggests that the changes incurred by OBSL1 are somehow adaptive, however these conservative amino acid changes do not seem to impact function strongly, as neither structure nor binding potential seem to be affected. It is possible that these conservative changes, when combined, have an effect that could not be discovered using these simple analyses, but which are somehow adaptive. After all, certain non-conservative mutations in OBSL1 have been shown to be highly deleterious in humans, causing pathologies such as the primordial growth disorder 3-M syndrome (Huber et al., 2010; Hanson et al., 2012). Such deleterious mutations would no doubt be removed through purifying selection in the harsh conditions of the deep sea.

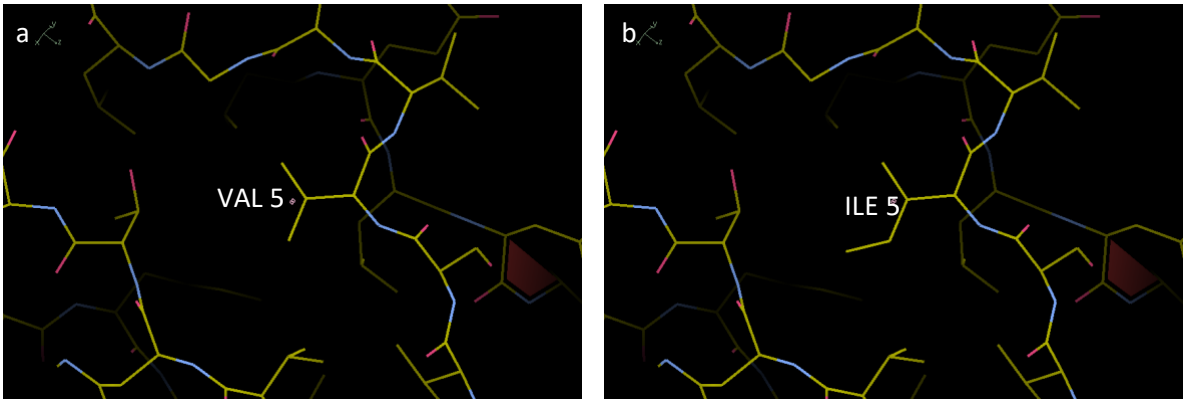


Figure 22: Zoomed in portion of the domain shown in Figure 21. SNP 11 causes valine (a) to be substituted for isoleucine (b): a very conservative change.

It is also possible that what seem to be non-impactful mutations actually have functions in the genome that we cannot see at this point. SNPs have been shown to influence alternative gene splicing, and therefore protein function (Hull et al., 2007). Even synonymous SNPs have been shown to influence messenger RNA splicing, stability and structure as well as protein folding (Hunt et al., 2009). It may be, therefore, that the main role of the SNPs that appear to produce conservative amino acid changes might lie in altering transcription and processes associated with messenger RNA. Whether this is the case, or whether the function of these conservative changes is more of a ‘sum of the whole’ impact, cannot be ascertained with these methods. Further study is required to reveal the function of these SNPs.

One key trend observed across the OBSL1 SNPs was that in multiple cases, the amino acid substitution resulted in tighter intra-domain binding in the shallow-living individuals than in those from deeper habitats. Though these changes do not affect the structure of the Ig domains, it is highly possible that they affect their movement and function. In a steered molecular dynamics simulation, Lu et al. (1998) demonstrated that the similar titin Ig domains unfolded and refolded again in coordination with muscle contraction, serving as molecular springs. Though the role of titin Ig domains is slightly different to that of OBSL1, with the former providing elasticity parallel to the muscle filaments, and the latter providing shock absorption and stability in the perpendicularly aligned M-band, the principles at play may be quite similar.

In titin, unfolding of Ig domains requires a single force peak to break the backbone hydrogen bonds connecting the A and G, and A' and G strands (Lu et al., 1998). Elastic refolding may then occur spontaneously, with H-bonds between the A' and G strands reforming rapidly (Gao et al., 2001). It is of particular note, therefore, that five SNPs: 5, 7, 15, 17, and 21, all affect H-bonds between the A and G, or A' and G strands (see Table 10). All of these SNPs cause substitutions from an amino acid which forms H-bonds between sidechains on those strands in the shallow form, to residues which do not permit such binding in the deep form. Two of the substitutions are even found on the same domain (Figure 23). The studies of titin Ig domain unfolding focus on backbone H-bonds, while the *C. rupestris* OBSL1 SNPs cause changes in sidechain interactions. Nevertheless, the positioning of the affected amino acids can be considered in the context of elasticity and folding, as they affect the binding between the same protein strands. It is clear that Ig domain unfolding requires greater force in the shallower living fish as a result of H-bonds between the A and G, and A' and G strands.

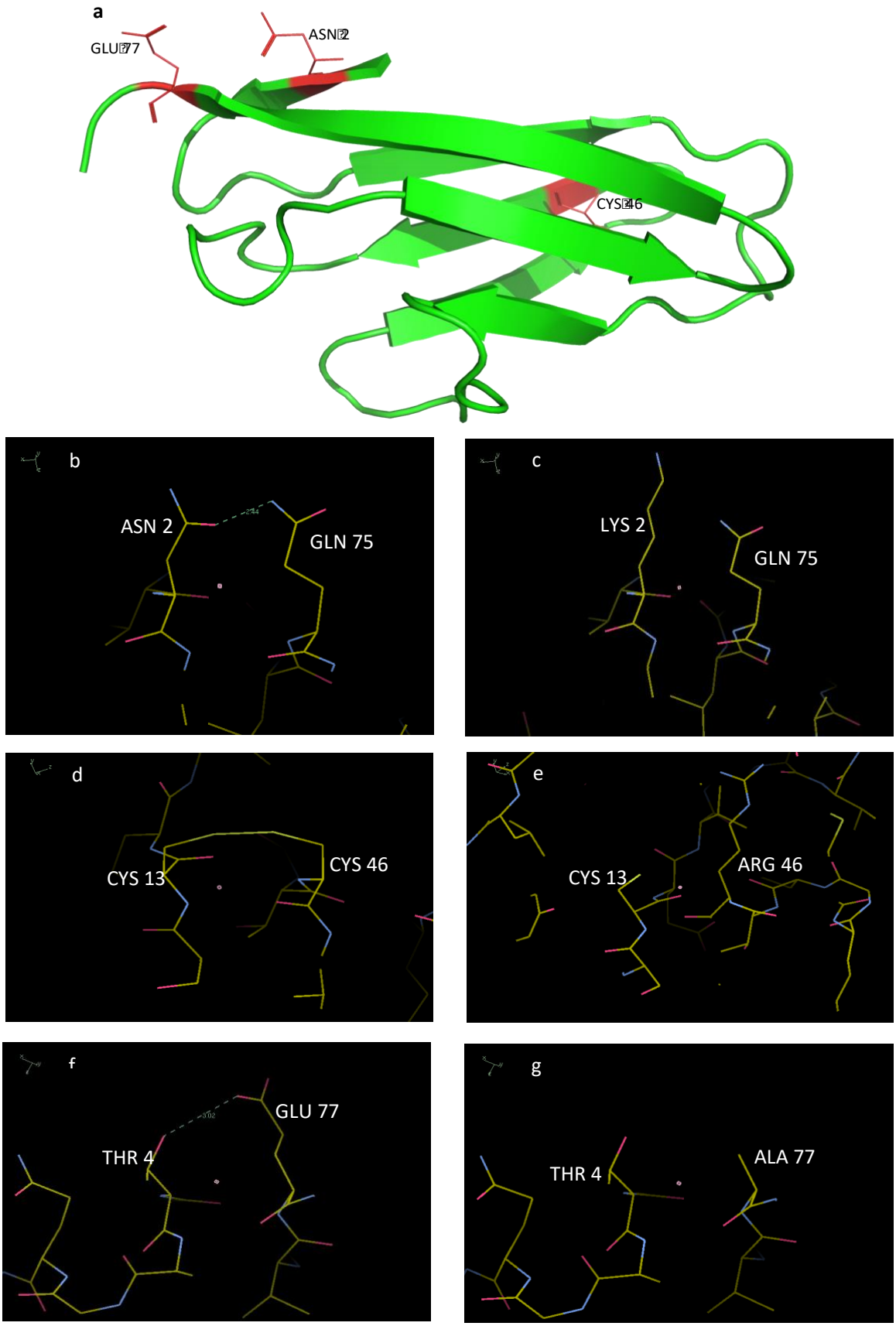


Figure 23: Representation of an Ig domain affected by SNPs 5, 6, and 7. (a) Model of the whole domain. SNP 5 causes asparagine (ASN 2) to change to lysine; SNP 6 causes cysteine (CYS 46) to change to arginine; and SNP 7 causes glutamic acid (GLU 77) to change to alanine. Affected amino acids are highlighted in red with wire structures showing the position of sidechains. (b) Wild type form of SNP 5 permits asparagine to form a hydrogen bond with a nearby glutamine, strengthening the interactions between strands A' and G. (c) Deep habitat mutant form of SNP 5 converts

asparagine to lysine, removing potential for hydrogen bonding with the G strand. (d) Wild type form of SNP 6 allows a disulphide bond to form between two cysteine residues across the Ig sandwich. (e) Deep habitat mutant form of SNP 6 converts cysteine to arginine, removing potential for disulphide bonds. (f) Wild type form of SNP 7 permits a hydrogen bond to form between glutamic acid and threonine, strengthening the interactions between strands A' and G. (g) Deep habitat mutant form of SNP 7 converts glutamic acid to alanine, removing potential for hydrogen bonding.

There are further examples of substitutions which appear to reduce the strength of intra-domain bonding in fish from deeper habitats. Though the binding in question does not bridge the A and G, or A' and G strands, and is therefore unlikely to affect protein unfolding as strongly, the trend still exists, and may act to reinforce the strength of Ig domains for the shallow living fish. For example, SNP 6 affects the same domain which displays two of the aforementioned substitutions affecting the A/A' and G strands (Figure 23). SNP 6 results in a change from cysteine, which can form a strong disulphide bridge across the middle of the Ig sandwich, to arginine, which can form no such bond, and might additionally disrupt the interactions within the Ig sandwich, as arginine is a particularly long and polar residue. This change from extremely strong internal binding in the Ig domain to a lack of such interactions, and potential disruption of the domain in fish from deeper habitats, supports the idea that intra-domain binding is stronger in the shallower living fish. Additional SNPs which cause tighter intra-domain binding in fish from shallower habitats included SNPs 9, 13, and 23 (see Table 10). This consistent trend for increased intra-domain binding in the shallow living fish suggests that these changes are adaptive, and should be considered in terms of potential selection pressures.

These findings could be associated with behaviour. The shallow habitat protein form may be better suited to the function of shock absorption suggested by Lange et al. (2019), as it can experience greater level forces without unravelling. If this is the case, the adaptation is likely to be in response to differences in muscle use by shallow and deep living fish. In Chapter 2, I outlined how bathypelagic *C. rupestris* individuals were smaller and more elongated than their mesopelagic conspecifics, and suggested that this difference in body form might have arisen from a selection for more efficient swimming in the bathypelagic. If

this is accepted as true, the improved shock absorption of OBSL1 in mesopelagic fish might be seen as a response to the more strenuous muscle use experienced with reduced swimming efficiency.

Alternatively, this trend could be linked to food availability. The characteristics of the bathypelagic might cause *C. rupestris* individuals to engage in different locomotory and predatory behaviours than they would use in the mesopelagic. Fairly little is known about the predatory behaviours of *C. rupestris*, though Drazen and Sutton (2017) classed them as demersal micronektonivores, feeding on small midwater fishes and cephalopods, engaging in active swimming. It is possible that in the bathypelagic, individuals engage in less active pursuit of prey items, as visibility is poorer, and energy conservation is more important (Warrant et al., 2003). These changes would result in lesser strain through prolonged and vigorous activity on the skeletal muscles of the bathypelagic fish, thereby permitting less effective shock absorption by OBSL1 in the M-band.

Hydrostatic pressure was expected to play some role in OBSL1 adaptation to depth. In particular, an increase in rigidity was anticipated in the proteins from the bathypelagic fish. No such trend was found, however, and this may be because of the intrinsic structural properties of Ig domains. Somkuti et al. (2013) analysed the denaturation behaviour of one titin Ig domain under a variety of different pressures and temperatures, and showed that, at 30°C, the domain was denatured by a pressure of 10.5 kbar, equivalent to 10362.7 atm. Increasing the temperature reduced the pressure at which the domain was denatured, meaning that at the low temperatures of the deep sea, an even greater pressure would be required to denature such a domain. To reach such a pressure naturally in the deep sea, depth would need to exceed 10000 m: well outside the 2600 m depth range of *C. rupestris* (Cohen, 1990). Though additional forces are present in living systems as a result of elastic stretching during muscle use which might impact OBSL1 denaturing, it seems unlikely that the relatively small hydrostatic pressure gradient experienced by *C. rupestris* would have much

effect. Perhaps the structure of Ig domains gives them natural resistance to denaturation at high hydrostatic pressure, or maybe the depth range of *C. rupestris* merely does not exert high enough hydrostatic pressures to impinge on Ig domain function.

Overall, it does seem that there could be some form of selection at play across the depth range of *C. rupestris*, which is impacting OBSL1. Though no known binding sites have been affected by the observed SNPs, changes in outwards-facing amino acids may impact interactions with other proteins or cellular components in a way that could not be analysed fully here. Conservative amino acid changes could add together to perform an unknown function, or the SNPs responsible for those changes may themselves have some function in altering gene splicing and transcription. Finally, the trend for less tight intra-domain binding in fish from deeper habitats can be linked to the lesser demands on the muscles of those individuals as a result of either decreased swimming activity or increased swimming efficiency from more elongated body shapes. It is clear that, though these findings build effectively on those of Gaither et al. (2018), greater study is required in this area to elucidate the whole host of functions that the OBSL1 SNPs perform. The obvious next step would be to attempt to synthesise the protein or a set of Ig domains, and to analyse its dynamics under changing temperature and pressure conditions, as well as investigating its elasticity, for both the shallow and deep form proteins. Such experiments would help to clarify the main selection pressures responsible for the changes in OBSL1 with habitat depth.

# Chapter 4: Beyond *C. rupestris*: Changes to Protein Structure and Function Across the Abyssal Boundary in *Coryphaenoides*

## Introduction

*Coryphaenoides* is a genus which spans a large variety of habitats and depth ranges. Recently, Gaither et al. (2016a) produced a phylogeny demonstrating the breadth of this variation and the relationships between species. This phylogeny demonstrated that abyssal *Coryphaenoides* species arose secondarily to those from shallower habitats, though any specific adaptations to facilitate this shift are currently unknown. It is therefore important to reach beyond our current understanding of the genetic markers of *Coryphaenoides* speciation by analysing features such as protein structure and function. Here, I will analyse the muscle protein OBSL1, which underwent changes associated with habitat depth across the meso-bathypelagic boundary in *C. rupestris*, as discussed in Chapter 3. This will provide valuable insight into the adaptations necessary for meso- and bathypelagic species to invade the abyss, and the different selection pressures at work across the pelagic zones.

*The abyssal zone*

Though not all authorities agree on the precise boundaries of the abyssopelagic or abyssal zone, a common definition is that it occupies the region between 3000 and 6000 m depth in the ocean (Priede, 2017). This is the largest depth zone in the ocean, accounting for 53.2% of the Earth's surface. The specific differences between the bathypelagic and abyssopelagic are somewhat ill-defined, partially owing to the logistical difficulties of generating an adequate data set to describe the conditions of the abyss (Angel, 1997). What is certain is that environmental conditions in the abyss are even more harsh than those in the bathypelagic. With the last of the sun's light being filtered out in the mesopelagic, the abyss is completely dark (Priede et al., 2017). Hydrostatic pressure continues to increase at ~100 kPa with every 10 m of depth (Childress, 1995), necessitating specific adaptations. Food

availability is reduced even further than in the bathypelagic. The only source of energy at these depths comes from the sinking of organic materials, yet according to Buesseler et al. (2007), as little as 0.5% of net primary productivity from the epipelagic may penetrate beyond 2000 m depth. In addition, overall reductions in faunal biomass in the abyss mean prey items for predatory species are scarce, and many abyssal mobile piscivores must include scavenging in their feeding repertoires to get enough food (Drazen and Sutton, 2017). The absence of certain taxa from the abyss such as rays and sharks has been linked to the significant energetic constraints imposed by this environment (Priede et al., 2006).

*Study protein and rationale*

Protein adaptations are necessary for survival in the abyss because of the energetic constraints and extremely high hydrostatic pressures. Although temperature is similarly low across the bathypelagic and abyssal zones, not resulting in much of a gradient for selection (Emery and Dewar, 1982), the consistent increase in hydrostatic pressure with increasing depth creates a stark selection pressure. This means that we might expect proteins to show increased rigidity and structural stability in the abyss compared to the bathypelagic, as these characteristics have been linked to the ability to withstand high pressures (Morita, 2003; Somero, 1998; Wakai et al., 2014; Porter et al., 2016).

After OBSL1 was shown to undergo multiple intrinsic changes which could be linked to adaptation to the bathypelagic in *C. rupestris* (Chapter 3), the question arose as to whether this protein underwent further changes in species adapted to the abyssopelagic. A muscle protein thought to undergo adaptations associated with behavioural and locomotory changes across the meso-bathypelagic boundary, OBSL1 represented a good study protein to explore adaptation across the bathy-abyssopelagic boundary.

The *Coryphaenoides* genus is, conveniently, perfect for exploring this question further, as it includes upwards of eight abyssal species thought to have arisen secondarily from their meso- and bathypelagic congeners (Gaither et al., 2016a). An ongoing project is taking place

which seeks to explore the genomic changes which have occurred in the divergence of the abyssal lineage, and I was able to use these data to build upon my own analyses. This project has involved sequencing the genomes of 14 *Coryphaenoides* species: 7 abyssal, and 7 non-abyssal (see Table 11 for a list of the species included). These genomes have been aligned against the annotated *C. rupestris* genome created by Gaither et al. (2018), and I was able to use the loci identified in the previous chapter to compare changes across these 14 species at the abyssal boundary.

Table 11: *Coryphaenoides* species with sampled genomes which were compared in this project. The genomes were aligned against the *C. rupestris* genome generated by Gaither et al. (2018). Depth ranges and pelagic zone affinities are given.

| **Species** | **Depth Range (m)** | **Abyssal or Non-Abyssal?** |
|---|---|---|
| ***C. acrolepis*** | 300 - 3700 | Non-abyssal |
| ***C. cinereus*** | 150 - 3500 | Non-abyssal |
| ***C. filifer*** | 1200 - 3000 | Non-abyssal |
| ***C. guentheri*** | 800 - 2900 | Non-abyssal |
| ***C. mexicanus*** | 100 - 1600 | Non-abyssal |
| ***C. rudis*** | 600 - 2400 | Non-abyssal |
| ***C. rupestris*** | 180 - 2600 | Non-abyssal |
| ***C. armatus*** | 300 - 5200 | Abyssal |
| ***C. brevibarbis*** | 1500 - 4700 | Abyssal |
| ***C. carapinus*** | 400 - 5600 | Abyssal |
| ***C. leptolepis*** | 600 - 4000 | Abyssal |
| ***C. mediterraneus*** | 1000 - 4300 | Abyssal |
| ***C. profundicolus*** | Unknown - 4900 | Abyssal |
| ***C. yaquinae*** | 3400 - 5800 | Abyssal |

OBSL1 was chosen for this comparison of abyssal vs. non-abyssal species because it is known to undergo change across the meso-bathypelagic boundary. The selection pressures are different across the meso-bathypelagic and bathy-abyssopelagic boundaries; factors such as light intensity do not change across the latter, where they create important gradients across the former. While I concluded in Chapter 3 that the changes observed in OBSL1 were unlikely to be associated with hydrostatic pressure, the much greater pressures of the abyssopelagic might be enough to necessitate a change in OBSL1 in response to pressure. As hydrostatic pressure has such a strong impact on protein function and adaptation

(Somero, 1998), I hypothesised that, in contrast to the meso-bathypelagic, hydrostatic pressure might represent a key selection gradient across the bathy-abyssopelagic boundary.

Proteins associated with OBSL1 in the M-band, titin and myomesin, were analysed in a similar way to explore whether any concurrent adaptations were occurring in those proteins along with OBSL1. This analysis will give a more holistic understanding of how different proteins are adapted to different habitat depths, and also shed more light on the evolutionary dynamics being experienced by *C. rupestris* across its own depth range.

## Methods

As the reference genomes for the 13 *Coryphaenoides* species analysed in addition to *C. rupestris* had already been created, data files were retrieved listing the SNPs which were found across those genomes compared to the *C. rupestris* genome. SNPs in the gene regions of titin, myomesin and OBSL1 were all collected. The process for analysing the effects of SNPs in OBSL1 was very similar to that described in Chapter 3. Species were ordered in the data sheet according to whether they live in the abyssal environment. Genotype ambiguity codes at SNP sites were colour coded, and this made it possible to identify any SNPs which were fixed at the abyssal boundary. Fixed SNPs from coding regions were extracted, and then tested for whether they were synonymous. This process was the same as that described in the Chapter 3, where the gene sequence for both the abyssal and non-abyssal forms were translated using ExPASY (Gasteiger et al., 2003), and then run through the EMBOSS Water pairwise alignment tool (Madeira et al., 2019). This allowed identification of non-synonymous SNPs. The domains of the non-synonymous SNPs were run through homology modelling program ModWeb, and the resultant structures analysed with Coot and PyMol.

The same process was carried out for titin and myomesin. Fgenesh (Solovyev et al., 2006) was used to identify coding regions in both proteins. SNPs across the whole myomesin gene sequence were analysed for fixation at the abyssal level and non-synonymity in the same way as described above. Titin, as the longest filament protein in the body (Zou et al., 2006),

had a gene sequence which was too long to effectively analyse in this way; the total count of SNPs across the titin gene region came to 6422. As a result, I chose to analyse only the M10 domain, which binds to OBSL1 in the M-band. The human titin M10 domain protein sequence was taken from PDBe (entry 2y9r; Pernigo et al., 2010), and the gene sequence was obtained by running the M10 protein sequence through NCBI protein BLAST (Madden, 2002). The M10 gene sequence was then found by carrying out a pairwise alignment between the whole titin gene sequence and the M10 protein sequence using the GeneWise protein-gene pair alignment tool (Madeira et al., 2019). The M10 gene sequence could then be run through the NCBI BLAST tool against the *C. rupestris* genome, which retrieved the *C. rupestris* gene sequence for the M10 domain. When this gene sequence was aligned against the gene sequence for the whole contig on which titin was found using the EMBOSS Water pairwise alignment tool (Madeira et al., 2019), it was possible to identify the location of the gene sequence on the contig, and therefore to find the relevant SNPs that affected it. The same process as described above was then used to identify non-synonymous fixed SNPs and analyse their impact on protein function.

One further analysis was carried out which was intended to explore the relative mutational load on each of the three proteins: OBSL1, titin and myomesin. Data files were generated with the overall number of synonymous and non-synonymous SNPs across the whole of the contigs of the proteins in question. The SNPs from the gene sequences of the proteins themselves were isolated, and the number of synonymous and non-synonymous SNPs counted. Percentages were calculated for the total number of nucleotides which were affected by SNPs across the length of the protein gene sequence, as well as the number of those SNPs which were non-synonymous.

## Results

Across the whole length of the contig on which OBSL1 coding regions were found, 106 SNPs were found which were differentiated at the abyssal boundary. This fixation was very clear, with different homozygous genotypes above and below the abyssal boundary (see Figure 24a). Of these 106 fixed SNPs, 46 were both found on coding regions and non-synonymous. Homology models could be generated for 39 of these 46 domains, including one which showed an amino acid change which was the result of two SNPs only one nucleotide apart. Key themes in this analysis were an apparent increase in rigidity in the abyssal proteins, mostly caused by replacements of glycine residues, whose minimal sidechains allow the protein backbone a lot of flexibility (SNPs 1, 7, 16, 17, 19 and 37 in Table 12). Many changes appeared to affect the external interface of the protein, but their impacts could not be fully understood with our incomplete knowledge of binding partners. Finally, though some SNPs appeared to cause some level of tighter intra-domain binding in shallower species analogous to that shown in the *C. rupestris* analyses (Such as SNPs 6, 20, 23, 30, and 34 in Table 12), this trend was not nearly as prominent, and none of the interactions strengthened the binding between the A/A' and G strands, which were most affected in that previous analysis. A full description of the impacts of the SNPs on OBSL1 function is displayed in Table 12.

When the same analyses were carried out on myomesin, fixation was much less clear than in OBSL1 (see Figure 24b); there was not such a clear segregation of abyssal and non-abyssal genotypes. Nonetheless, 46 SNPs could roughly be categorised as fixed at the abyssal boundary. Of these, however, only four SNPs were both found on coding regions and non-synonymous for the myomesin protein. Furthermore, the changes incurred by these SNPs seemed to have much more conservative impacts on function than those found in OBSL1 (for a full description of changes, see Table 13).

**a** **OBSL1**

| Species | SNP ID | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| *C. acrolepis* | GG | AA | GG | CC | GG |
| *C. cinereus* | GG | AA | GG | CC | GG |
| *C. filifer* | GG | AA | GG | CC | GG |
| *C. guentheri* | GG | AA | GG | CC | GG |
| *C. mexicanus* | GG | AA | GG | CC | GG |
| *C. rudis* | GG | AA | GG | CC | GG |
| *C. rupestris* | GG | AA | GG | CC | GG |
| *C. armatus* | AA | TT | AA | GG | TT |
| *C. brevibarbis* | AA | TT | AA | GG | TT |
| *C. carapinus* | AA | TT | AA | GG | TT |
| *C. leptolepis* | AA | TT | AA | GG | TT |
| *C. mediterraneus* | AA | TT | AA | GG | TT |
| *C. profundicolus* | AA | TT | AA | GG | TG |
| *C. yaquinae* | AA | TT | AA | GG | TT |

**b** **Myomesin**

| Species | SNP ID | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| *C. acrolepis* | TT | AA | AA | TT |
| *C. cinereus* | TT | AA | AA | TT |
| *C. filifer* | TT | AA | AA | TT |
| *C. guentheri* | TT | AA | AA | TT |
| *C. mexicanus* | TT | AA | AA | TT |
| *C. rudis* | TT | AA | AA | TT |
| *C. rupestris* | TT | AA | AA | TT |
| *C. armatus* | GG | CC | AA | CC |
| *C. brevibarbis* | GG | AA | GG | CC |
| *C. carapinus* | N | AA | GG | CC |
| *C. leptolepis* | GG | CC | GG | CC |
| *C. mediterraneus* | GG | CC | GG | CC |
| *C. profundicolus* | GG | CC | GG | CC |
| *C. yaquinae* | GG | N | AA | CC |

**c** **Titin**

| Species | SNP ID |
|---|---|
| | 1 |
| *C. acrolepis* | GG |
| *C. cinereus* | GG |
| *C. filifer* | GG |
| *C. guentheri* | CC |
| *C. mexicanus* | CC |
| *C. rudis* | CC |
| *C. rupestris* | CC |
| *C. armatus* | GG |
| *C. brevibarbis* | GG |
| *C. carapinus* | AA |
| *C. leptolepis* | GG |
| *C. mediterraneus* | GG |
| *C. profundicolus* | GG |
| *C. yaquinae* | GG |

Figure 24: SNP nucleotide identities for each of the 14 species analysed. Both nucleotides of the diploid pair are represented. Colour codes are assigned according to nucleotide identity. The non-abyssal species are listed first, separated by a thicker black line from the abyssal species. a) OBSL1 nucleotide identities for the first five SNPs modelled. Clear fixation can be observed at the abyssal boundary for all SNPs. b) Myomesin nucleotide identities for all SNPs modelled. Fixation is a little weaker than in OBSL1. c) Nucleotide identities for the only non-synonymous SNP within the M10 domain. Fixation at the abyssal boundary is very weak.

Table 12: List of amino acid changes and their effects from the non-synonymous SNPs found to be fixed at the abyssal boundary across the coding regions of the OBSL1 gene. SNPs are numbered according to the order in which they affect the protein, with 1 at the N-terminal end. Table continues on pages 90 and 91.

| SNP | Non-abyssal amino acid | Abyssal amino acid | Analysis |
|---|---|---|---|
| **1** | Glycine | Aspartic acid | Affected amino acid is at linker region between beta strands. Glycine has no sidechain, but aspartic acid has a carboxylic acid sidechain which faces outwards from the Ig domain. Aspartic acid carries a negative charge which is a large change from the non-charged glycine residue, and adds to a region of negative charge as other aspartic acid residues are present nearby. This might strengthen inter-protein binding. This SNP affects Ig1, which binds to titin M10, however the affected amino acid does not appear to affect the binding surface for this interaction. Glycine allows greater protein chain flexibility, and the abyssal change to aspartic acid would create a more rigid structure. |
| **2** | Lysine | Asparagine | Affected amino acid is found on a linker region between Ig domains. Lysine has a long sidechain ending in an amide group capable of forming ionic bonds with positive ligands. Asparagine does not have the potential for such bonds. The sidechains of these amino acids face outwards from the Ig domain. |
| **3** | Valine | Isoleucine | Affected amino acid is found on a linker region between beta strands. Both residues are aliphatic and hydrophobic. The sidechains from these residues face inwards v domain. Valine has a larger, less flexible sidechain. The hydrophobic core may therefore be tighter and less disrupted in the abyssal form. Same domain as SNP 4. |
| **4** | Glutamine | Glutamate | Amino acid sidechain faces outwards from the Ig domain. Glutamine has a flexible sidechain ending in an amide group, giving it a high potential for H bonding. No H bond recipients were close enough to form such a bond, however. Glutamate has a long, flexible sidechain with a carboxylic acid group, giving it a strong negative charge suitable for ionic bonding. Again, however, no suitable binding partners were found. Same domain as SNP 5. |
| **5** | Valine | Leucine | SNP affects amino acid on linker region between beta strands. This change is very conservative, with one simple aliphatic residue changing to another. |
| **6** | Threonine | Alanine | Threonine may form a H bond with nearby 26 LEU, forming a tighter bond between the linker regions next to the B and E strands. Alanine is aliphatic and is not able to perform such bonding. |
| **7** | Glycine | Arginine | Affected amino acid faces outwards from the domain. The change from glycine would increase rigidity in the abyssal form. Arginine has a long, positively charged sidechain which does not appear to interact with any of the structures on the Ig domain itself, though it joins a large number of similarly large, positively charged residues on this surface of the protein. Increasing this area of positive charge may play a role in external interactions. Same domain as SNPs 8 and 9 (Figure 25). |
| **8** | Leucine | Valine | Leucine and valine have largely similar properties, and both sidechains face outwards from the Ig domain. Same domain as SNPs 7 and 9 (Figure 25). |
| **9** | Lysine | Glutamine | The affected amino acid faces outwards from the domain. The change from lysine to glutamine changes the sidechain charge from positive to polar. Lysine can form ionic bonds where glutamine cannot. This change might affect any interface with other proteins or cell components but no binding partners are known. Same domain as SNPs 7 and 8 (Figure 25). |
| **10** | Isoleucine | Threonine | The amino acid is next to a beta hairpin; isoleucine faces outwards and does not appear to impact the structure of the protein. The introduction of threonine creates the potential for hydrogen bonding, however no suitable binding partners are found on the domain itself. Instead, this change could affect binding to external ligands. |

Table 12 continued.

| | | | |
|---|---|---|---|
| **11** | Lysine | Arginine | The same substitution, in the same place on the domain, as SNPs 13 and 14. Lysine has long sidechain with positive charge at the end, and arginine is similar but has a greater positive charge due to presence of a guanidine group. It is possible that the long positive sidechains which face outwards are responsible for some kind of extra-protein interaction, which may be stronger in the abyssal variants due to the greater charge of arginine. Same domain as SNP 12. |
| **12** | Glutamic acid | Aspartic acid | Affected amino acid is on a complex linker region. Both residues have carboxylic acid groups, and the substitution does not appear to change much. Same domain as SNP 11. |
| **13** | Lysine | Arginine | The same substitution, in the same place on the domain, as SNPs 11 and 14. The same effects are likely. |
| **14** | Lysine | Arginine | The same substitution, in the same place on the domain, as SNPs 11 and 13. The same effects are likely. Same domain as SNP 15. |
| **15** | Arginine | Lysine | Essentially the opposite change to SNPs 11, 13 and 14. May balance out the change of SNP 14 as the two SNPs are on the same domain, or may impact external interactions by lessening the potential for interactions in the abyssal form. |
| **16** | Glycine | Serine | Removal of glycine results in increased rigidity of the protein in the abyssal form. Residue is on a hairpin structure. Serine can form H bonds unlike glycine, but does not appear to interact with any residues within the Ig domain. Same domain as SNPs 17 and 18. |
| **17** | Glycine | Valine | Removal of glycine results in increased rigidity of the protein in the abyssal form. Residue is on a hairpin structure. Valine is hydrophobic but is modelled as facing away from the domain; it is more likely that this residue would face into the hydrophobic core. Cannot therefore draw full conclusions on its interactions. Same domain as SNPs 16 and 18. |
| **18** | Threonine | Serine | A conservative change which does not appear to affect any binding or structural elements. Same domain as SNPs 16 and 17. |
| **19** | Glycine | Cysteine | Removal of glycine results in increased rigidity of the protein in the abyssal form. Residue is on a hairpin, and the cysteine sidechain faces outwards. This change may impact external binding. |
| **20** | Serine | Cysteine | Affected residue faces outwards from the E strand. Serine may form a H bond with the nearby 52 GLU on the D strand, but cysteine does not have this capability. |
| **21** | Threonine | Alanine | Threonine may form a H bond with the nearby 60 THR, strengthening the binding between the B and E chains. Alanine cannot form such a bond. . Same domain as SNPs 22 and 23. |
| **22** | Arginine | Lysine | Amino acid is on a linker region. Arginine may form H bonds with both 30 ASP and 28 TRY, both of which are also on linker regions, however lysine can bind with only one of these residues at a time. It is unclear whether such bonds actually occur in vivo, however. Same domain as SNPs 21 and 23. |
| **23** | Threonine | Isoleucine | Threonine on the F strand may form a H bond with 83 ALA (G strand). Such bonding cannot occur when the aliphatic isoleucine is substituted. Same domain as SNPs 21 and 22. |
| **24** | Serine | Asparagine | Potential for H bonding is the same in these two residues, and no significant impact is had by the change. |
| **25** | Arginine | Glutamine | Amino acid is on a linker region between beta strands, facing outwards. No intra-domain interactions can be seen, so change may affect external binding. Arginine can form ionic bonds, which glutamine cannot. Same domain as SNP 26. |

Table 12 continued.

| | | | |
|---|---|---|---|
| **26** | Methionine | Leucine | Amino acid faces outwards from beta strand. It is unclear what impact this change may have on external binding. |
| **27** | Lysine | Asparagine | The change from the long, negatively charged lysine sidechain to the polar asparagine sidechain removes the potential for ionic bonds with the externally-facing residue, however no binding partners are known. |
| **28** | Lysine | Isoleucine | Negatively charged lysine is replaced by non-charged isoleucine, changing the protein surface charge and potentially affecting external binding. Again though, no binding partners are known for this domain. |
| **29** | Serine | Alanine | No intra-domain bonding seems to occur with the sidechain of either residue, though serine has the potential to form H bonds where alanine does not. Same domain as SNPs 30 and 31. |
| **30** | Asparagine | Histidine | Asparagine on the B strand may form a H bond with the aromatic region of nearby 85 TRP on the E strand, strengthening the binding between those strands. Histidine cannot perform such an interaction. Same domain as SNPs 29 and 31. |
| **31** | Methionine | Threonine | The methionine residue is on the D strand and faces inwards towards the hydrophobic domain core, but does not appear to interact with any other residues. When threonine is substituted in, it can form a H bond with nearby 85 TRP on the E strand. This may strengthen the intra-domain binding in the abyssal form. Same domain as SNPs 29 and 30. |
| **32** | Alanine | Proline | Alanine has a simple aliphatic sidechain, where proline's sidechain curves around, binding with the backbone N atom. This greatly reduces backbone flexibility in the abyssal form. Same domain as SNPs 33 and 34. |
| **33** | Glutamine | Asparagine | Note that two SNPs affecting nucleotides spaced with only one nucleotide between them caused this change. The affected amino acid sidechain faces outwards from the domain, and both glutamine and asparagine have sidechains ending in amide groups. This change looks quite conservative. In *C. yaquinae* only, the amino acid changes to lysine, which carries a positive charge and would change the protein surface quite drastically if any binding were to occur. Same domain as SNPs 32 and 34. |
| **34** | Asparagine | Serine | Asparagine on the F strand may form a H bond with the G strand 86 THR. Though serine could technically form such a bond, the distance between the residues becomes too great as its sidechain is shorter. Same domain as SNPs 32 and 33. |
| **35** | Aspartic acid | Alanine | Negatively charged aspartic acid is capable of forming ionic and H bonds. Alanine does not have this capacity. The aspartic acid residue is part of a region of negative charge on the domain surface along with 15 GLU. The substitution may reduce external binding potential. Same domain as SNPs 36, 37 and 38. |
| **36** | Glutamine | Alanine | Glutamine has a long carboxylic acid sidechain, where alanine has a simple aliphatic sidechain. This change may affect binding potential of the domain. Same domain as SNPs 35, 37 and 38. |
| **37** | Glycine | Serine | Glycine is substituted for serine, increasing the rigidity of the backbone structure. Same domain as SNPs 35, 36 and 38, and amino acid is adjacent to that affected by SNP 38. |
| **38** | Lysine | Threonine | Amino acid faces outwards from the domain, and the change from lysine to threonine represents a change from a positive to a polar sidechain. No intra-domain binding could be found. Any impact this SNP might have would be related to external interactions. Same domain as SNPs 35, 36 and 37, and amino acid is adjacent to that affected by SNP 37. |
| **39** | Aspartic acid | Tyrosine | The sidechain faces outwards from the domain. Aspartic acid has a negatively charged carboxylic acid sidechain, where tyrosine has a polar aromatic sidechain. This is a drastic change, though its function is unclear at this stage. |

Table 13: List of amino acid changes and their effects from the non-synonymous SNPs found to be fixed at the abyssal boundary across the coding regions of the myomesin gene. SNPs are numbered according to the order in which they affect the protein, with 1 at the N-terminal end.

| SNP | Non-abyssal amino acid | Abyssal amino acid | Analysis |
|---|---|---|---|
| **1** | Glycine | Alanine | Fairly conservative change in terms of sidechain function, however the substitution of glycine would result in lesser backbone flexibility in the abyssal form. |
| **2** | Leucine | Valine | Both amino acids are simple and aliphatic, so could be quite conservative. It is found on the binding surface which interacts with OBSL1, however; the slightly shorter sidechain of valine may alter the binding site slightly. |
| **3** | Asparagine | Serine | Sidechain of the affected amino acid faces outwards and doesn't appear to interact with anything within the domain, and no binding partners are known. Both residues are polar, and the change appears to be quite conservative. Same domain as SNP 4. |
| **4** | Valine | Alanine | A change between two simple aliphatic residues. This change is conservative. |

When the titin M10 domain was analysed, five SNPs were found in coding regions which were fixed at the abyssal boundary. Only one of those SNPs was non-synonymous, however, and it showed very weak fixation (see Figure 24c). The resultant amino acid substitution is on a surface of the domain far away from the binding surface which interacts with OBSL1. The SNP causes a substitution from lysine to asparagine. Both sidechains face away from the domain, so if this SNP has an effect it will be in external interactions with other proteins or cell components. The poor fixation of the SNP and the lack of knowledge of any further binding partners for this domain, however, mean that no strong effect or adaptation can be linked to this change.

The final analysis for these three muscle proteins concerned the abundance of synonymous and non-synonymous SNPs. Though this analysis did not include fixation data, it still represented the frequency and type of changes which occurred in the gene sequences of these three proteins across the abyssal boundary. SNPs were labelled as 'synonymous', 'missense' meaning non-synonymous, and 'intron' meaning the SNP is found on an intron and so has no synonymous or non-synonymous identity. As can be seen in Table 14, OBSL1 and titin had a similar proportion of nucleotides affected by SNPs, where myomesin had around half

as many. Although more of titin's SNPs were found in coding regions than in OBSL1, OBSL1 had a much greater proportion of non-synonymous to synonymous SNPs, and indeed was the only protein to have more non-synonymous than synonymous SNPs. Myomesin had a particularly low proportion of non-synonymous SNPs.

Table 14: Breakdown of SNP identities across the whole of the OBSL1, titin and myomesin gene regions. '% Gene Region Including SNPs' indicates the proportion of nucleotides across the whole gene affected by SNPs. The other % values represent the proportion of those SNPs of each type: synonymous, non-synonymous and intron. The ratio column refers to the relative number of non-synonymous to synonymous SNPs, where a value above 1 indicates a greater frequency of non-synonymous SNPs.

| **Gene** | **% Gene Region Including SNPs** | **% Synonymous SNPs** | **% Non-Synonymous SNPs** | **% Intron SNPs** | **Ratio Non-Synonymous to Synonymous SNPs** |
|---|---|---|---|---|---|
| **OBSL1** | 4.91 | 13.4 | 15 | 71.6 | 1.12 |
| **Titin** | 4.74 | 27.2 | 20.6 | 50.6 | 0.76 |
| **Myomesin** | 2.17 | 28.3 | 10.5 | 61.1 | 0.37 |

## Discussion

Overall, the changes observed do appear to align with the hypothesis that increasing hydrostatic pressure creates a key selection gradient across the bathy-abyssopelagic boundary. One of the biggest trends across all the SNPs was the substitution of glycine in the non-abyssal form to be replaced by another amino acid in the abyss. As glycine has only a hydrogen atom as its sidechain, it permits a great deal of flexibility to the protein backbone. Further, the lack of any hydrogen bonds or hydrophobic interactions from this residue decreases overall stability of the protein (Jagodzinski et al., 2012). It would therefore appear that abyssal OBSL1 is more rigid and stable than its bathypelagic homologues. These changes in rigidity and stability are especially pronounced when the glycine residue in question is found on a flexible region like a hinge (Akbal-Delibas et al., 2013). In this analysis of OBSL1, a total of six SNPs caused substitutions of a non-abyssal glycine residue to another amino acid in the abyss. In four of those cases, the glycine residue is found on a flexible linker or hairpin region between beta strands. It is therefore likely that changes

which increase rigidity of these regions in particular have an even greater effect in increasing protein rigidity in the abyssal form. This finding is in line with the expectations of protein depth adaptation from the literature: the increasing hydrostatic pressure associated with abyssal habitats requires increased rigidity for proteins to withstand. One further piece of evidence suggesting a selection for more rigid proteins in the abyss is the insertion of a proline residue by SNP 32. Proline is the most rigid amino acid, as its sidechain curves around to bind to the backbone nitrogen atom. This restricts the conformational freedom of the protein backbone, effectively increasing overall rigidity. Substitutions to proline have been linked to increased rigidity in the context of improved thermostability (Hardy et al., 1993), and this goes hand in hand with improved resistance to pressure (Somero, 1998).

As OBSL1 is a structural protein responsible for shock absorption in the sarcomeric M-band, flexibility and elasticity are crucial to its function (Lange et al., 2019). OBSL1 appears to be undergoing an increase in rigidity in the abyssal *Coryphaenoides* species analysed here, and this change could be associated with increasing hydrostatic pressure in deeper habitats. What is not so obvious, however, is what impact this change might have on the function of the protein itself. Though the changes caused by these SNPs could theoretically permit the protein to survive without denaturation in the abyss, it may come at a functional cost. There are examples of this occurring in other protein types such as the dehydrogenases (Somero and Seibenaller, 1979), where resistance to high hydrostatic pressure is conferred at the cost of functional efficiency. OBSL1 flexibility and elasticity appears to be important for its function of stabilising the sarcomere, so increased rigidity may reduce its performance. This change may in turn be associated with an alteration in muscle use by abyssal species. It may be that abyssal species compensate for less effective stabilisation of the sarcomere by using their muscles less intensely. This is hard to demonstrate, however, owing to the difficulties of conducting behavioural studies in the deep sea. Overall, it seems that OBSL1 is undergoing an increase in rigidity in response to the high hydrostatic pressures of the abyss.

This change may come at the cost of function in deeper habitats, although further analyses are required to establish this.

15 of the SNPs in OBSL1 caused changes in amino acids whose sidechains faced outwards from the Ig domain with no discernible impact on the function of the domain itself. Some of these changes were quite substantial; for example, SNP 7 caused a change from the small, uncharged glycine to the long and positively charged arginine (Figure 25). Though this change does not affect any known binding site, such a dramatic change may have an impact on interactions which are as of yet unknown. As discussed previously, not all the complex associations of the M-band are understood at present (Lange et al., 2019), and interactions with unknown proteins, or other cellular components, is possible. With the present analysis, however, it is impossible to ascertain the role that these externally-facing changes might play in adaptation to the abyss, as potential binding partners are unknown.

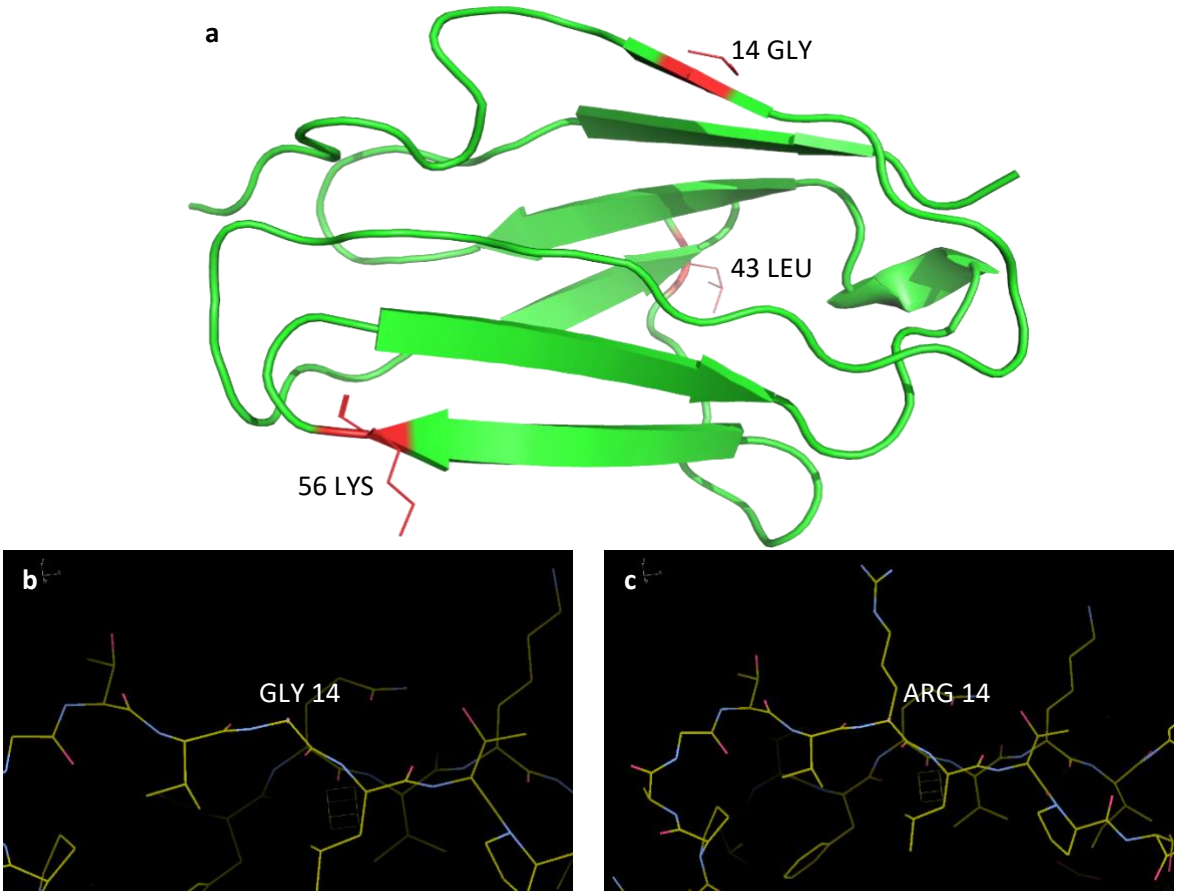


Figure 25: a) Model representing the OBSL1 domain affected by SNPs 7, 8 and 9. Amino acids affected by SNPs are highlighted in red, and their sidechains represented by wire structures. b) Zoomed in image of GLY 14, which is affected by SNP 7. This is the non-abyssal form of the protein. c) Zoomed in image of ARG 14, which is found in the abyssal form of this protein. The striking difference in sidechain size and binding potential can be observed when comparing b) and c).

When *C. rupestris* was analysed across the meso-bathypelagic boundary (Chapter 3), the clearest trend caused by the SNPs was an increase in intra-domain binding between amino acid sidechains in the mesopelagic. This was particularly notable between the A/A' and G strands, which are most central to the elastic function of Ig domains (Lu et al., 1998; Gao et al., 2001). Though some analogous changes are observed here, the trend is much less pronounced across the bathy-abyssopelagic boundary. SNPs 6, 20, 23, 30, and 34 affect amino acids which can form H bonds with other nearby residues in the bathypelagic, but not the abyssal form. None of these bonds, however, link the A/A' and G strands as was repeatedly observed in Chapter 3. The conclusion that can be drawn from this finding is that the selection pressures acting on OBSL1 are not the same across the meso-bathypelagic boundary as across the bathy-abyssopelagic boundary.

Overall, it appears that the key selection pressure acting on OBSL1 at the abyssal boundary is hydrostatic pressure. This is demonstrated by the consistent removal of glycine residues in abyssal species, as well as the addition of proline, which increases backbone rigidity and decreases pressure sensitivity. Though five changes analogous to those in *C. rupestris* which decreased potential for intra-domain binding in deeper living individuals were observed, these changes were far less impactful than those in Chapter 3, and the overall impact is not comparable. This further supports the role of hydrostatic pressure as a selection pressure for OBSL1 in the abyss which is not experienced in shallower ocean zones.

Titin and myomesin were also analysed in order to gauge whether any similar adaptations were occurring. Overall, the changes observed in these proteins were far less substantial than those of OBSL1. First, the fixation of SNPs in these proteins was not nearly as clear as in OBSL1 (see Figure 24), suggesting that any selection is not acting as strongly across the abyssal boundary in titin and myomesin as in OBSL1. Second, the proportion of synonymous and non-synonymous SNPs in each of these proteins demonstrated an increased trend for non-synonymous change in OBSL1 compared to the others. Though titin had a similar

overall proportion of SNPs across its gene region as OBSL1, the proportion of those SNPs that were non-synonymous was 5% lower. Myomesin had less than half the total proportion of SNPs as the other two proteins. The proportion of non-synonymous SNPs was also around half of that observed in OBSL1. Finally, the non-synonymous SNPs themselves were extremely conservative in myomesin compared to OBSL1. These factors could suggest a much lower mutation rate, or a greater level of purifying selection, in this protein.

Though one change from glycine in the bathypelagic to another residue in the abyss suggested an increase in rigidity in the abyssal form of myomesin, the other three SNPs observed in this protein had minimal effect. Though titin is too large to carry out these same analyses on, the one non-synonymous SNP found on the M10 domain did not appear to have any clear adaptive role either. The modelling analyses seem to support the conclusion that less significant change is occurring in titin and myomesin compared to OBSL1. These are very simple metrics by which to analyse selection and adaptation, and they should not be considered definitive. It may be that titin and myomesin have fewer non-synonymous mutations because they are undergoing a greater level of purifying selection, where potentially deleterious mutations are selected against. This cannot be determined without a more in-depth analysis of dN/dS ratios, however. Alternatively, these results could suggest that OBSL1 is undergoing a greater level of mutational change than either titin or myomesin. Considering the apparent adaptive function of many of these changes in OBSL1, it may be tentatively suggested that OBSL1 is undergoing more intense selection than the other two muscle proteins analysed here.

In conclusion, it appears that hydrostatic pressure is acting as a selection pressure for OBSL1 across the bathy-abyssopelagic boundary in *Coryphaenoides*. The changes observed were consistent with expectations for increased protein rigidity with increasing habitat depth. Furthermore, it appears that OBSL1 undergoes more substantial change at this boundary than other M-band proteins titin and myomesin. Why OBSL1 seems to be demonstrating

greater adaptation to increasing habitat depth than other proteins with similar structures and functions is unclear. All three proteins analysed have elastic functions permitted by their repeating Ig domain structure, and are responsible for maintaining order in muscle filaments. Further exploration of these proteins is required to understand why OBSL1 appears to be undergoing greater change across the bathy-abyssopelagic boundary.

# Chapter 5: Synthesis and Conclusion

The findings of this project support the hypotheses made by Gather et al. (2018), who stated that genotypic segregation across the meso-bathypelagic boundary suggested the existence of two distinct ecotypes of *C. rupestris*, despite apparent random mating throughout habitat depths. Here, I have demonstrated some of the morphological and protein related differences observed in these meso- and bathypelagic ecotypes. I have also contrasted these findings with data from abyssal and non-abyssal congeners of *C. rupestris*, demonstrating the changes in relative importance of different selection pressures in the mesopelagic, bathypelagic, and abyssopelagic. It is now important to look at these findings collectively, in the context of the existing literature, to appreciate their relevance to our overall understanding of adaptation within the deep sea ecosystem.

*Relevance of ecotype segregation to conservation*

Adaptation and ecotype segregation according to depth in *C. rupestris* must be considered in the context of fishery and conservation. Currently endangered in Europe and critically endangered globally, the *C. rupestris* stock was greatly depleted by fishing in the 1970s and is currently in decline (Koslow et al., 2000; Iwamoto, 2015; Cook et al., 2015). *C. rupestris* is a slow-growing species with highly irregular recruitment, and the impact of historical overfishing is likely still felt today (Koslow et al., 2000; Baker et al., 2009). Historically, fishing effort has been concentrated shallower than 1500 m (Atkinson, 1995; Neat and Burns, 2010). It is therefore relevant that the ecotype segregation I have found here also takes place at around 1500 m, with all juveniles and gravid females found shallower than this depth. Fishing at these depths would disproportionately remove juveniles and reproducing individuals, further damaging the population's opportunity to stabilise. Furthermore, such practices would result in an exaggerated removal of the shallow ecotype compared to the deep. As the two ecotypes are thought to engage in random mating (Gaither et al., 2018), this may have an unknown impact on the population as a whole. It may result

in more deep ecotype fish being born, either skewing the species' preferred habitat depth to be deeper over several generations, or resulting in a population crash from overpopulation in the bathypelagic, where food resources may be lower.

While directed fishery of *C. rupestris* has ceased, fisheries do still remove individuals through bycatch. This is arguably more dangerous, as regulations on bycatch are more difficult to implement. The impact of bycatch on *C. rupestris* is severe, particularly from the Greenland halibut fishery. Data from the 1990s estimated that bycatch rates for *C. rupestris* were the highest for any species in that fishery (Devine and Haedrich, 2008). More recently, bycatch estimates for *C. rupestris* were as high as 1,000 tonnes per year (Simpson et al., 2011). It is likely that bycatch, similarly to directed fishing, selectively removes *C. rupestris* individuals from shallower habitats at a disproportionate rate; halibut fishery, the main source of bycatch, is also concentrated shallower than the 1500 m level (Bowering and Nedreaas, 2000).

It is clear that limiting fishery to areas above a certain depth is not sufficient for protecting species like *C. rupestris*, whose habitats extend across a large depth range. Though theoretically, targeted fishery shallower than 1500 m would provide a refuge in deeper waters for this species, I have shown that the localisation of juveniles and gravid females to shallower habitats where fishing takes place is the real concern. It is essential that any future conservation regulations such as Marine Protected Areas or closed areas should take this into account, and act across greater depth ranges. If no such regulation is established, the impact on the recovery of the population may be even more severe than currently predicted (Simpson et al., 2011), as the role of this ecotype segregation by habitat depth was not previously considered.

*Linking morphology with protein function: the role of locomotion*

Several of the findings of the morphological analysis of *C. rupestris* linked ecotype differentiation in the bathypelagic to a decrease in food availability (Chapter 2). Food

limitation places selection pressures on energy conservation and storage as well as feeding ecology. Indeed, all key morphological findings from Chapter 2 could be linked in some way to improving efficiency of energy acquisition, storage and use in this energy poor environment. Swimming efficiency in bathypelagic individuals was thought to be improved with smaller, more slender bodies and greater lipid storage in the swimbladder to improve buoyancy (Pelster, 1997; Neat and Campbell, 2013). Greater relative eye and gape sizes were associated with deeper habitats, improving feeding opportunities (Drazen and Sutton, 2017). Liver lipid storage also increased in deeper waters, potentially reducing starvation risk when regular meals cannot be counted on (Musick and Cotton, 2015). It is possible that this key selection pressure also impacts behaviour and foraging strategies, and that the findings from the muscle protein analysis might build on our understanding of how adaptation to food limitation or changing food resources might occur at increasing habitat depths.

There is a recognised trend for lower metabolic rates, as well as a reduction in diel vertical migration, in bathypelagic species compared to mesopelagic, with energy conservation as the driving selection pressure (Childress et al., 1980; Childress, 1995). If this change results in decreased locomotory activity in deeper habitats, the body muscles would experience less strain. As a result, OBSL1 and other M-band proteins might face lesser requirements in shock absorption. This is in line with the protein analysis of meso- and bathypelagic *C. rupestris* in Chapter 3. Tighter intra-domain bonding was shown in Ig domains across the length of OBSL1 in the mesopelagic fish compared to bathypelagic. This could be linked to the greater elastic demand placed on the protein in mesopelagic fish due to higher activity levels. The consistent reinforcements made to the A/A’ to G strand bonds in mesopelagic fish support this idea, as the separation of these strands is particularly relevant to the elastic function of the Ig domain (Lu et al., 1998; Gao et al., 2001). Energy conservation is a selection pressure which may therefore explain both my morphological findings and my protein analyses. Lowered metabolic rate goes hand in hand with this change, potentially

reducing locomotory activity in the bathypelagic and relaxing the demand for strong elastic action of OBSL1.

Light availability also appears to impact locomotion. The visual interactions hypothesis suggests that decreasing light levels such as those experienced across the meso-bathypelagic boundary reduce the distances over which predator-prey interactions take place, thereby relaxing the selection pressure for locomotory capacity (Salvanes and Kristoffersen, 2001; Drazen and Siebel, 2007). The larger relative eye sizes of bathypelagic *C. rupestris* suggests that the darkness of the bathypelagic does have an impact on ecotype adaptation. If light availability is a factor which impacts eye morphology, it may well also impact behaviour and locomotion as suggested by Salvanes and Kristoffersen (2001). Fewer bouts of quick swimming to catch prey or avoid predators in the bathypelagic would reduce the need for such strong elastic action in the OBSL1 Ig domains in a similar way to that described above. Overall, locomotory activity may be the crucial link between the findings of the morphological and protein analyses which I have performed. Morphological markers for greater energy conservation and a reduction in muscle protein elasticity are associated with selection pressures for low activity levels in the bathypelagic.

When the protein analyses from the abyssal and non-abyssal *Coryphaenoides* species are analysed, they do not fit in with the lines of argument outlined above. This is because the environmental changes across the bathy-abyssopelagic are not the same as those across the meso-bathypelagic boundary. Light intensity, for example, remains at zero when moving from the bathypelagic into the abyss. Food availability remains limited; in both the bathy- and abyssopelagic, sinking organic particles are the sole energy source (with the exception of hydrothermal vents; Grabowski et al., 2019). Because of this, OBSL1 should not have been expected to follow the same depth-related trends across the bathy-abyssopelagic boundary as the meso-bathypelagic. Instead, new selection pressures arose: namely, hydrostatic pressure. SNPs repeatedly caused changes which increased protein rigidity in the

abyss compared to the bathypelagic. Similar changes have been well-documented as responses to the increased hydrostatic pressure experienced at greater depths, for example in LDH and opsin proteins (Somero, 1998; Brindley et al., 2008; Porter et al., 2016).

The protein changes observed in the abyssal species must also be considered in the context of overall lifestyle and behaviour. If metabolic rate is expected to continue to decrease as species' depth ranges grow deeper, a reduction in locomotion in abyssal species would be seen in turn which may then be linked to the changes seen in OBSL1. There is little evidence that metabolic rate does decrease neatly as habitat range deepens in this genus, however. Siebenaller et al. (1982) demonstrated that LDH activity, an indicator of a fish's capacity for high energy burst swimming, varied amongst five *Coryphaenoides* species independently of depth, and more likely changed according to interspecific variation in feeding habits. This suggests that, where feeding mode and energy conservation play important roles in *C. rupestris* adaptation across the meso-bathypelagic boundary, these factors do not impact OBSL1 in its adaptation to the abyss. Instead, it appears that hydrostatic pressure was the main selection pressure which impacted this protein during *Coryphaenoides'* invasion of the abyss.

*Further considerations*

The intrinsic protein changes observed in OBSL1 have demonstrated some useful findings concerning the selection pressures experienced in different ocean depth zones. These changes alone, however, do not tell the whole story. Adaptations to protein function can be implemented in a variety of ways, including through external changes. For example, osmolytes found in teleost muscle like betaine and trimethylamine oxide (TMAO) stabilise proteins, allowing them to continue functioning under high hydrostatic pressure. Concentrations of these osmolytes increase with habitat depth, with abyssopelagic species demonstrating much higher levels than bathypelagic and mesopelagic species in turn (Yancey et al., 2001; 2004). Increases in muscle osmolyte concentration may add to the

intrinsic increases in structural stability observed in abyssal species in this project. Further to this, an extrinsic increase in osmolyte levels in bathypelagic *C. rupestris* compared to mesopelagic may compensate for the apparent lack of any intrinsic changes in OBSL1 which increase pressure resistance in the bathypelagic form. Exploration of how the mesopelagic, bathypelagic and abyssal forms of OBSL1 behave in different osmolyte concentrations would be useful to fully elucidate this relationship.

Though I have been able to link adaptations of OBSL1 in the meso-, bathy-, and abyssopelagic to set of selection pressures, the role of other proteins in ecotype and species segregation according to depth remain somewhat unclear. First, there is the question of why the other muscle proteins I investigated, titin and myomesin, appear to undergo fewer changes at increasing habitat depths than OBSL1. These proteins both have similar structures and localities to OBSL1, and myomesin performs a very similar function in the M-band (Agarkova and Perriard, 2005; Pernigo et al., 2010). It is therefore unclear why they appear to have undergone such comparatively small changes with depth in both the single species analysis of *C. rupestris*, and the study of the 14 abyssal and non-abyssal species. As fewer non-synonymous SNPs were found in these proteins, continuing to analyse them using these in silico methods is unlikely to yield solid results. If any adaptive changes to titin and myomesin do occur with depth, they may be extrinsic, related to processes such as alternate splicing or alteration of osmolyte concentration. These possibilities should be explored if improved understanding of the adaptive dynamics of this set of proteins is to be improved.

*Conclusion*

At the beginning of this project, I set out to explore the phenotypic impacts of the genotypic segregation of *C. rupestris* according to depth. This was with the aim to elucidate key selection pressures and mechanisms for adaptation across a depth gradient in the deep sea. I believe I have done this, demonstrating how changes to morphology and protein function are associated with different selection pressures in the meso- and bathypelagic, and

extending those analyses to contrast with findings about abyssopelagic species. Though further study is required to fully understand the adaptive processes at play, this project gives a good insight into the population dynamics of *C. rupestris* individually, and into the most crucial selection pressures for adaptation to different habitat depths in the deep sea.

# Appendices

Appendix 1: Raw morphometric data for each of the 139 *C. rupestris* specimens sampled, identified by unique numbers (ID No.). Measurements were taken three times each, and the mean displayed here. Measures are pre-anal fin length (PAFL), head length (HL), head depth (HD), pre-orbital length (POL), inter-orbital width (IOW), left orbit diameter (ODL), body depth (BD), maximum body width (MW), gape size (GS), and gape length (GL).

| ID No. | PAFL | HL | HD | POL | IOW | ODL | BD | MW | GS | GL |
|---|---|---|---|---|---|---|---|---|---|---|
| **01026** | 130.20 | 94.30 | 68.90 | 23.47 | 32.67 | 28.30 | 74.60 | 35.30 | 59.17 | 33.43 |
| **01027** | 94.90 | 67.77 | 50.37 | 16.37 | 25.87 | 23.80 | 54.00 | 26.93 | 32.77 | 22.83 |
| **01028** | 134.00 | 98.43 | 71.97 | 25.60 | 39.33 | 30.37 | 83.30 | 41.40 | 54.53 | 33.90 |
| **01029** | 147.00 | 100.33 | 77.47 | 27.23 | 40.30 | 32.67 | 83.53 | 43.70 | | 33.07 |
| **01030** | 121.00 | 88.70 | 68.57 | 25.93 | 30.67 | 26.30 | 75.90 | 32.03 | 56.10 | 30.40 |
| **01031** | 123.33 | 88.93 | 67.93 | 24.40 | 31.33 | 28.33 | 72.73 | 32.40 | 55.47 | 32.30 |
| **01032** | 131.00 | 95.50 | 83.10 | 25.40 | 39.37 | 31.30 | 85.00 | 43.53 | 63.10 | 37.23 |
| **01033** | 132.00 | 93.43 | 73.50 | 19.97 | 37.50 | 31.13 | 79.87 | 39.60 | 57.93 | 35.63 |
| **01034** | 113.97 | 87.13 | 64.40 | 20.53 | 32.33 | 29.13 | 69.47 | 31.90 | 50.47 | 29.10 |
| **01035** | 131.07 | 88.77 | 65.50 | 24.30 | 33.40 | 25.57 | 72.30 | 36.70 | 56.57 | 28.80 |
| **01036** | 107.70 | 79.20 | 61.40 | 17.97 | 30.50 | 26.07 | 65.63 | 26.50 | 50.37 | 28.07 |
| **01037** | 95.00 | 64.83 | 52.53 | 15.20 | 24.57 | 22.03 | 57.53 | 25.23 | 42.50 | 23.67 |
| **01038** | 151.00 | 105.67 | 83.20 | 27.23 | 43.37 | 32.33 | 85.70 | 42.53 | 64.53 | 36.10 |
| **01039** | 136.00 | 105.53 | 80.40 | 24.40 | 42.23 | 34.03 | 89.20 | 37.17 | 62.50 | 38.97 |
| **01040** | 192.00 | 137.33 | 107.67 | 31.33 | 57.00 | 37.00 | 121.67 | 59.33 | 36.00 | 45.67 |
| **01041** | 228.00 | 152.23 | 121.63 | 44.07 | 62.80 | 41.27 | 126.73 | 82.00 | 51.27 | 48.07 |
| **01042** | 160.00 | 111.70 | 83.47 | 25.47 | 46.17 | 33.27 | 93.27 | 48.00 | 62.63 | 35.23 |
| **01043** | 125.00 | 97.03 | 76.60 | 24.10 | 37.30 | 29.37 | 76.23 | 34.50 | 41.87 | 24.83 |
| **01044** | 164.00 | 122.67 | 95.70 | 28.27 | 46.50 | 38.63 | 96.90 | 43.00 | 66.17 | 36.77 |
| **01045** | 169.50 | 123.33 | 93.33 | 33.53 | 51.90 | 35.07 | 95.90 | 52.23 | 49.23 | 47.83 |
| **01046** | 80.85 | 62.67 | 46.73 | 16.90 | 26.10 | 19.43 | 47.93 | 22.00 | 34.17 | 18.33 |
| **01047** | 135.10 | 96.47 | 67.00 | 22.77 | 37.63 | 29.30 | 73.37 | 30.67 | 53.80 | 28.77 |
| **01048** | 159.50 | 114.27 | 85.20 | 28.33 | 42.67 | 34.97 | 93.60 | 39.00 | 63.57 | 36.43 |
| **01049** | 185.50 | 137.67 | 105.03 | 35.90 | 51.43 | 38.00 | 117.37 | 57.37 | 41.30 | 43.60 |
| **01050** | 156.00 | 112.93 | 95.13 | 24.77 | 51.17 | 33.90 | 97.37 | 48.67 | 67.53 | 33.73 |
| **01051** | 126.80 | 96.30 | 70.67 | 24.43 | 36.63 | 30.07 | 77.47 | 39.00 | 45.60 | 31.00 |
| **01052** | 156.75 | 115.77 | 104.67 | 27.80 | 46.67 | 31.83 | 109.90 | 49.67 | 63.90 | 38.73 |
| **01053** | 146.60 | 111.40 | 82.40 | 27.00 | 44.67 | 33.33 | 86.17 | 35.17 | 60.70 | 33.60 |
| **01054** | 154.00 | 111.43 | 85.67 | 26.10 | 43.20 | 32.97 | 87.73 | 42.00 | 63.57 | 29.43 |
| **01055** | 197.50 | 142.07 | 113.63 | 31.67 | 56.30 | 42.10 | 113.07 | 68.00 | 81.60 | 43.83 |
| **01056** | 144.60 | 109.73 | 90.73 | 24.33 | 46.97 | 33.90 | 101.90 | 43.17 | 66.07 | 33.43 |
| **01057** | 139.45 | 94.27 | 72.40 | 21.20 | 35.60 | 28.60 | 76.37 | 32.00 | 48.50 | 29.03 |
| **01058** | 146.65 | 103.60 | 81.63 | 25.37 | 42.43 | 30.90 | 89.37 | 46.67 | 50.37 | 31.87 |
| **01059** | 140.75 | 106.10 | 89.37 | 26.73 | 45.73 | 32.53 | 86.50 | 45.33 | 62.00 | 33.80 |
| **01060** | 84.00 | 64.00 | 54.47 | 14.37 | 24.40 | 20.13 | 50.47 | 20.33 | 32.37 | 18.27 |
| **01061** | 92.75 | 71.40 | 54.53 | 21.20 | 26.37 | 22.30 | 54.50 | 28.57 | 39.90 | 20.03 |
| **01062** | 104.10 | 75.37 | 56.57 | 18.43 | 26.70 | 24.27 | 57.37 | 30.10 | 36.60 | 22.30 |
| **01063** | 116.45 | 84.37 | 65.77 | 20.37 | 31.30 | 26.10 | 65.93 | 32.50 | 44.80 | 22.67 |
| **01064** | 78.70 | 58.50 | 46.27 | 13.60 | 22.37 | 17.40 | 39.57 | 16.00 | 31.53 | 17.53 |
| **01065** | 57.10 | 46.00 | 32.90 | 12.47 | 13.97 | 16.00 | 31.53 | 14.07 | 24.40 | 13.30 |
| **01066** | 59.40 | 42.90 | 30.60 | 12.73 | 13.50 | 14.37 | 32.43 | 13.47 | 24.13 | 12.77 |
| **01067** | 53.00 | 40.50 | 28.57 | 10.60 | 15.00 | 12.67 | 29.60 | 13.33 | 19.57 | 14.83 |
| **01068** | 60.15 | 42.90 | 29.60 | 12.77 | 12.70 | 13.27 | 29.03 | 14.07 | 20.90 | 13.90 |
| **01101** | 218.00 | 145.00 | 118.00 | 31.70 | 56.10 | 41.17 | 121.00 | 68.33 | 86.17 | 49.67 |
| **01102** | 195.00 | 133.00 | 105.00 | 32.87 | 51.20 | 41.10 | 113.00 | 62.33 | 77.03 | 45.67 |
| **01103** | 190.00 | 125.67 | 103.00 | 29.87 | 47.47 | 40.13 | 117.00 | 47.00 | 78.00 | 46.13 |
| **01105** | 201.00 | 140.00 | 113.00 | 34.60 | 55.60 | 39.63 | 121.00 | 60.00 | 89.47 | 48.37 |
| **01106** | 188.00 | 131.33 | 101.67 | 33.63 | 52.20 | 37.63 | 106.43 | 53.00 | 73.57 | 43.73 |
| **01107** | 193.33 | 129.67 | 106.00 | 29.20 | 53.23 | 40.27 | 118.00 | 57.00 | 73.47 | 48.00 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **01108** | 172.00 | 119.33 | 96.30 | 25.07 | 44.90 | 41.63 | 97.47 | 46.67 | 70.57 | 40.40 |
| **01109** | 214.00 | 145.33 | 122.50 | 36.40 | 58.00 | 44.37 | 135.67 | 64.50 | 90.33 | 57.27 |
| **01110** | 188.00 | 128.33 | 98.63 | 33.17 | 51.60 | 38.43 | 113.40 | 59.00 | 74.97 | 44.07 |
| **01111** | 163.00 | 114.00 | 93.00 | 27.30 | 41.30 | 36.27 | 101.70 | 48.67 | 69.90 | 41.87 |
| **01112** | 169.00 | 110.83 | 99.00 | 25.70 | 41.40 | 39.73 | 110.60 | 52.00 | 69.20 | 42.90 |
| **01113** | 45.57 | 33.20 | 24.23 | 8.27 | 11.70 | 12.70 | 22.43 | 8.97 | 17.00 | 9.40 |
| **01114** | 75.20 | 57.10 | 42.97 | 12.77 | 21.17 | 21.83 | 44.07 | 18.30 | 33.90 | 19.40 |
| **01115** | 55.00 | 42.47 | 30.00 | 10.80 | 13.63 | 15.17 | 29.23 | 12.20 | 24.60 | 12.13 |
| **01116** | 58.13 | 45.73 | 35.07 | 12.13 | 18.03 | 16.03 | 31.97 | 16.03 | 28.90 | 15.13 |
| **01117** | 92.13 | 71.70 | 35.17 | 17.00 | 24.93 | 24.93 | 52.03 | 25.73 | 38.07 | 22.87 |
| **01118** | 96.90 | 69.70 | 52.50 | 18.53 | 25.37 | 22.83 | 53.90 | 22.00 | 41.70 | 23.50 |
| **01119** | 116.37 | 81.20 | 60.17 | 19.50 | 31.20 | 29.70 | 62.50 | 33.00 | 47.83 | 26.87 |
| **01120** | 118.90 | 88.20 | 63.60 | 20.97 | 31.33 | 29.87 | 67.50 | 38.00 | 52.40 | 28.97 |
| **01121** | 130.13 | 89.50 | 66.73 | 24.03 | 34.27 | 30.73 | 72.43 | 35.13 | 54.90 | 29.63 |
| **01122** | 120.17 | 88.50 | 64.50 | 22.43 | 30.83 | 27.33 | 69.47 | 37.10 | 47.70 | 29.30 |
| **01123** | 139.67 | 108.07 | 80.60 | 25.80 | 40.70 | 37.67 | 89.40 | 43.00 | 57.50 | 33.73 |
| **01124** | 128.23 | 93.50 | 70.50 | 23.23 | 35.33 | 31.10 | 72.73 | 38.50 | 58.60 | 33.23 |
| **01125** | 146.00 | 100.00 | 78.70 | 25.23 | 38.53 | 33.47 | 87.60 | 40.73 | 55.57 | 33.27 |
| **01126** | 152.83 | 109.13 | 83.17 | 27.33 | 42.30 | 32.73 | 90.00 | 49.00 | 63.30 | 35.97 |
| **01127** | 132.00 | 91.57 | 77.23 | 21.50 | 36.50 | 30.57 | 80.30 | 37.00 | 62.37 | 34.53 |
| **01128** | 155.17 | 102.27 | 82.10 | 22.87 | 40.33 | 30.10 | 87.83 | 43.50 | 66.40 | 31.17 |
| **01129** | 145.60 | 102.17 | 89.83 | 20.27 | 42.00 | 33.43 | 90.03 | 48.37 | 60.00 | 27.60 |
| **01130** | 143.20 | 100.23 | 76.03 | 26.80 | 40.33 | 30.97 | 83.37 | 46.17 | 62.90 | 34.73 |
| **01131** | 127.27 | 88.50 | 61.70 | 19.27 | 34.30 | 29.07 | 67.63 | 34.40 | 54.93 | 27.77 |
| **01132** | 137.40 | 96.97 | 77.23 | 21.87 | 38.50 | 31.97 | 81.60 | 36.83 | 61.27 | 34.87 |
| **01133** | 137.67 | 101.90 | 85.27 | 24.97 | 35.63 | 30.30 | 88.80 | 40.00 | 67.30 | 38.37 |
| **01134** | 167.33 | 111.87 | 95.37 | 29.30 | 47.07 | 35.57 | 100.67 | 46.83 | 73.50 | 37.83 |
| **01135** | 146.00 | 106.63 | 78.47 | 27.33 | 44.47 | 33.07 | 79.47 | 38.70 | 63.27 | 32.67 |
| **01136** | 123.70 | 93.93 | 77.23 | 25.17 | 35.07 | 29.77 | 74.17 | 39.00 | 60.77 | 33.50 |
| **01137** | 153.00 | 104.70 | 79.37 | 26.77 | 43.57 | 32.57 | 88.37 | 39.17 | 65.77 | 35.63 |
| **01138** | 184.00 | 121.73 | 97.27 | 28.67 | 45.50 | 34.87 | 108.07 | 52.83 | 71.40 | 42.37 |
| **01139** | 143.67 | 104.10 | 85.10 | 23.57 | 40.03 | 33.13 | 92.07 | 45.57 | 61.07 | 32.90 |
| **01140** | 129.67 | 92.23 | 68.43 | 19.37 | 39.07 | 28.23 | 75.33 | 36.20 | 59.90 | 27.77 |
| **01141** | 132.77 | 95.10 | 71.00 | 26.23 | 37.07 | 29.13 | 74.20 | 43.00 | 55.47 | 31.33 |
| **01142** | 150.00 | 105.67 | 83.90 | 24.10 | 44.80 | 33.67 | 98.03 | 38.90 | 60.90 | 35.03 |
| **01143** | 123.90 | 89.87 | 67.03 | 21.37 | 35.47 | 27.90 | 72.27 | 30.00 | 56.70 | 30.67 |
| **01476** | 182.00 | 130.00 | 106.00 | 30.60 | 49.90 | 35.13 | 105.50 | 55.00 | 71.70 | 35.30 |
| **01477** | 184.00 | 133.00 | 99.33 | 35.60 | 53.53 | 38.50 | 109.50 | 55.50 | 71.10 | 44.70 |
| **01478** | 175.00 | 122.33 | 101.70 | 30.23 | 49.87 | 36.83 | 105.30 | 54.17 | 70.63 | 45.47 |
| **01700** | 198.67 | 130.33 | 103.00 | 33.67 | 49.37 | 36.63 | 111.47 | 52.33 | 80.67 | 44.83 |
| **01701** | 181.67 | 135.00 | 110.00 | 30.70 | 51.47 | 40.07 | 118.00 | 60.00 | 77.80 | 49.87 |
| **01702** | 176.00 | 119.00 | 97.50 | 32.23 | 45.53 | 35.23 | 103.00 | 57.00 | 71.60 | 38.87 |
| **01703** | 198.00 | 135.33 | 108.00 | 33.47 | 56.20 | 38.83 | 120.73 | 63.33 | 78.77 | 46.43 |
| **01704** | 114.00 | 83.37 | 66.70 | 20.03 | 31.03 | 29.07 | 70.90 | 32.20 | 49.40 | 28.37 |
| **01705** | 180.67 | 129.00 | 102.00 | 31.77 | 51.70 | 37.33 | 114.00 | 66.00 | 72.97 | 45.97 |
| **01706** | 164.00 | 109.50 | 86.30 | 27.07 | 38.20 | 33.43 | 91.90 | 45.53 | 61.63 | 37.03 |
| **01707** | 182.00 | 127.33 | 108.00 | 33.67 | 52.40 | 40.00 | 124.00 | 62.00 | 74.60 | 45.90 |
| **01708** | 207.00 | 137.33 | 114.00 | 35.67 | 54.57 | 42.23 | 126.33 | 71.33 | 77.23 | 51.17 |
| **01709** | 162.67 | 120.33 | 97.67 | 32.13 | 47.13 | 32.97 | 106.40 | 52.00 | 65.10 | 40.60 |
| **01710** | 204.00 | 132.67 | 116.00 | 29.10 | 57.80 | 39.40 | 121.33 | 78.00 | | |
| **01711** | 221.00 | 157.33 | 124.00 | 39.80 | 64.57 | 44.13 | 143.00 | 74.00 | 79.97 | 54.53 |
| **01712** | 208.67 | 137.00 | 114.33 | 35.53 | 55.33 | 38.67 | 122.00 | 60.00 | 83.73 | 49.63 |
| **01713** | 208.00 | 139.00 | 108.00 | 35.60 | 54.77 | 41.10 | 121.00 | 67.00 | 80.80 | 48.50 |
| **01714** | 192.67 | 139.67 | 110.00 | 37.80 | 56.73 | 39.03 | 116.00 | 60.00 | 79.50 | 48.43 |
| **01715** | 166.33 | 120.33 | 100.23 | 30.83 | 48.93 | 35.93 | 107.50 | 50.00 | 67.10 | 40.90 |
| **01716** | 194.50 | 140.00 | 116.00 | 34.30 | 55.10 | 43.50 | 128.00 | 64.00 | 74.77 | 49.73 |
| **01717** | 193.00 | 132.00 | 104.00 | 35.50 | 54.17 | 38.07 | 113.00 | 68.67 | 78.13 | 45.83 |
| **01718** | 145.00 | 105.90 | 77.47 | 27.27 | 39.60 | 33.33 | 86.30 | 45.30 | 61.73 | 34.83 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **01719** | 99.07 | 75.63 | 55.30 | 17.73 | 28.33 | 26.27 | 56.70 | 30.47 | 45.37 | 25.47 |
| **01720** | 194.00 | 142.67 | 115.00 | 39.60 | 57.80 | 36.93 | 123.00 | 68.00 | 74.07 | 46.13 |
| **01721** | 230.67 | 159.33 | 121.00 | 44.93 | 63.23 | 42.50 | 126.67 | 62.00 | 80.87 | 49.90 |
| **01722** | 186.00 | 134.33 | 102.00 | 31.27 | 49.80 | 38.50 | 117.00 | 60.67 | 86.17 | 49.67 |
| **01723** | 177.00 | 110.33 | 107.67 | 19.87 | 47.83 | 35.03 | 111.67 | 64.00 | 71.47 | 38.80 |
| **01724** | 179.00 | 126.83 | 99.17 | 33.83 | 51.37 | 35.47 | 115.67 | 49.00 | 71.77 | 39.87 |
| **01725** | 208.00 | 141.00 | 111.00 | 37.17 | 56.30 | 37.83 | 124.00 | 57.00 | 80.50 | 45.43 |
| **01826** | 120.20 | 91.87 | 66.47 | 22.43 | 34.13 | 28.90 | 73.03 | 31.03 | 53.17 | 29.67 |
| **01827** | 116.40 | 85.70 | 63.27 | 19.53 | 31.20 | 26.87 | 71.13 | 32.27 | 53.17 | 30.00 |
| **01828** | 115.03 | 86.50 | 64.03 | 21.40 | 33.37 | 26.67 | 68.70 | 30.73 | 48.80 | 28.23 |
| **01829** | 137.73 | 96.40 | 79.63 | 21.07 | 37.00 | 30.83 | 83.07 | 34.80 | 54.87 | 28.93 |
| **01830** | 128.90 | 94.03 | 71.00 | 22.07 | 38.57 | 30.50 | 75.73 | 36.37 | 54.87 | 26.37 |
| **01831** | 127.40 | 95.43 | 71.70 | 25.23 | 35.03 | 29.30 | 75.43 | 35.00 | 56.23 | 32.30 |
| **01832** | 160.17 | 113.43 | 89.13 | 32.87 | 46.20 | 34.47 | 87.87 | 52.00 | 71.50 | 37.30 |
| **01833** | 149.33 | 104.97 | 80.73 | 26.93 | 40.47 | 33.23 | 86.23 | 41.00 | 66.37 | 40.77 |
| **01834** | 131.50 | 93.67 | 69.53 | 22.03 | 34.83 | 29.63 | 71.87 | 30.33 | 60.87 | 30.90 |
| **01835** | 151.00 | 107.20 | 84.43 | 24.33 | 39.80 | 37.23 | 97.00 | 39.00 | 64.67 | 36.77 |
| **01836** | 134.00 | 91.77 | 69.50 | 22.67 | 33.93 | 28.33 | 74.40 | 38.20 | 50.40 | 31.50 |
| **01837** | 165.00 | 114.40 | 81.60 | 26.30 | 45.07 | 34.93 | 97.37 | 42.00 | 67.60 | 42.30 |
| **01838** | 117.20 | 83.47 | 65.00 | 18.80 | 32.50 | 31.23 | 67.33 | 37.50 | 56.87 | 32.57 |
| **01839** | 144.00 | 100.63 | 78.50 | 24.80 | 35.93 | 31.60 | 81.10 | 39.30 | 62.90 | 37.67 |
| **01840** | 111.50 | 80.10 | 58.00 | 20.37 | 30.40 | 24.73 | 62.40 | 24.60 | 48.60 | 26.77 |
| **01841** | 135.00 | 100.70 | 74.80 | 25.70 | 42.77 | 34.27 | 85.80 | 39.80 | 59.90 | 36.07 |
| **01842** | 170.00 | 120.80 | 87.30 | 32.37 | 48.13 | 35.37 | 94.93 | 36.67 | 64.50 | 44.13 |
| **01843** | 155.00 | 117.13 | 96.00 | 30.43 | 46.03 | 36.90 | 94.40 | 51.00 | 62.20 | 37.53 |
| **01844** | 151.00 | 112.90 | 82.80 | 23.17 | 47.17 | 35.30 | 91.87 | 45.80 | 64.00 | 38.60 |
| **01845** | 135.00 | 96.37 | 73.00 | 28.00 | 34.43 | 26.47 | 73.10 | 43.17 | 55.07 | 34.60 |
| **01846** | 137.93 | 100.40 | 75.97 | 24.67 | 35.93 | 30.73 | 83.43 | 37.93 | 59.40 | 35.70 |
| **01847** | 147.00 | 99.27 | 72.30 | 27.03 | 41.80 | 33.73 | 68.30 | 46.00 | 56.33 | 36.00 |
| **01848** | 131.00 | 100.17 | 73.80 | 26.07 | 37.33 | 31.43 | 78.60 | 34.93 | 59.90 | 36.57 |
| **01849** | 133.00 | 99.00 | 75.40 | 23.63 | 41.57 | 33.30 | 78.20 | 34.70 | 64.80 | 36.10 |
| **01850** | 124.60 | 84.17 | 66.40 | 19.23 | 34.93 | 29.80 | 69.30 | 36.13 | 54.17 | 33.00 |

Appendix 2: Mass measurements from dissections of the 139 *C. rupestris* samples. Each fish is identified by a unique number (ID No.). Gill raker numbers from the upper and lower limb of the first arch are shown with notation 'upper limb+lower limb'. Liver, swimbladder, heart and gonad masses are also displayed. For some samples, heart mass is listed as '<1g'. These samples were collected during dissections without access to a high precision low mass scale, so precise masses could not be recorded.

| ID No. | Gill Rakers | Liver mass (g) | Swimbladder mass (g) | Heart mass (g) | Gonad mass (g) |
|---|---|---|---|---|---|
| **01026** | 3+15 | 18.5 | 1.6 | 0.338 | 1.147 |
| **01027** | 3+16 | 9.1 | 0.595 | 0.215 | 0.365 |
| **01028** | 3+16 | 37.2 | 2 | 0.48 | 1.6 |
| **01029** | 4+13 | 28.1 | 5.1 | 0.9 | 3.4 |
| **01030** | 3+15 | 14.5 | 4.4 | 0.42 | 1.7 |
| **01031** | 3+17 | 9.7 | 2.4 | 0.297 | 2.3 |
| **01032** | 3+16 | 19.4 | 8 | 0.6 | 3.9 |
| **01033** | 3+15 | 15.4 | 7.7 | 0.46 | 2.8 |
| **01034** | 3+15 | 19 | 1.1 | 0.236 | 1.2 |
| **01035** | 3+15 | 33.1 | 1.4 | 0.6 | 1.2 |
| **01036** | 3+15 | 4.9 | 1.1 | 0.235 | 0.22 |
| **01037** | 3+15 | 8.6 | 0.958 | 0.18 | 0.711 |
| **01038** | 3+15 | 17 | 4.6 | 0.6 | 3 |
| **01039** | 4+17 | 15.8 | 5.6 | 0.6 | 4.7 |
| **01040** | | 30.7 | 24.7 | 1.4 | |
| **01041** | | 50.5 | 57.3 | 1.8 | |
| **01042** | 4+14 | 48.8 | 6.2 | 0.9 | |

| | | | | | |
|---|---|---|---|---|---|
| **01043** | 4+14 | 13 | 1.9 | 0.3 | |
| **01044** | 4+15 | 28.8 | 7.9 | 0.8 | |
| **01045** | | 31.5 | 12 | 1.6 | |
| **01046** | 3+14 | 2.5 | 0.3 | 0.1 | |
| **01047** | 4+14 | 7.3 | 2.1 | 0.5 | |
| **01048** | 4+15 | 14.7 | 5 | 1.2 | |
| **01049** | | 48.8 | 19.1 | 2 | |
| **01050** | 3+16 | 19 | 6.5 | 0.8 | |
| **01051** | 3+15 | 30.7 | 1.7 | 0.8 | |
| **01052** | 3+16 | 12.1 | 7.2 | 0.9 | 3.1 |
| **01053** | 4+15 | 19.9 | 3.3 | 0.6 | 2.1 |
| **01054** | 3+13 | 12.5 | 2.2 | 0.878 | |
| **01055** | 4+15 | 35.5 | 17.7 | 3.2 | 11.4 |
| **01056** | 3+15 | 18.4 | 4.3 | 0.9 | 2.4 |
| **01057** | 3+14 | 18.3 | 1.4 | 0.59 | |
| **01058** | 4+14 | 36.3 | 1.3 | 0.8 | |
| **01059** | 4+15 | 22 | 3.3 | 0.9 | |
| **01060** | 4+15 | 2.6 | 0.2 | <0.1 | |
| **01061** | 4+14 | 6.5 | 0.44 | 0.328 | |
| **01062** | 4+15 | 6.7 | 0.663 | 0.22 | |
| **01063** | 3+14 | 7.2 | 1.2 | 0.313 | |
| **01064** | 4+15 | 1.3 | 0.195 | 0.077 | |
| **01065** | 4+14 | 1.3 | 0.07 | 0.04 | |
| **01066** | 4+14 | 1.1 | | 0.042 | |
| **01067** | | 1.1 | | <0.1 | |
| **01068** | 1+14 | 0.7 | 0.09 | 0.034 | |
| **01101** | 3+18 | 33.7 | 9.8 | 1.9 | 74.1 |
| **01102** | 3+15 | 18.2 | 6.8 | 0.9 | 29.1 |
| **01103** | 3+16 | 9.9 | 10.8 | 0.9 | 61.2 |
| **01105** | 3+16 | 57.6 | 3.1 | 1.2 | 1.9 |
| **01106** | 4+17 | 27.8 | 5.5 | 0.9 | 6.1 |
| **01107** | 3+18 | 20.5 | 7.1 | 1.7 | 59.6 |
| **01108** | 3+17 | 7.1 | 15.3 | 1 | 2.6 |
| **01109** | 3+16 | 21.5 | 23.9 | 1.5 | 86.3 |
| **01110** | 3+16 | 19.8 | 13 | 1.8 | 4.6 |
| **01111** | 4+16 | 11.9 | 11.2 | 0.658 | 1.756 |
| **01112** | 3+15 | 14.9 | 4.5 | 0.9 | 29.9 |
| **01113** | 3+13? | 0.253 | | 0.022 | |
| **01114** | 3+14 | 2.225 | 0.202 | 0.085 | |
| **01115** | 1+15? | | | 0.019 | |
| **01116** | 4+15 | 1.66 | | 0.074 | |
| **01117** | 3+15 | 3.32 | 0.549 | 0.144 | |
| **01118** | 3+14 | 4 | 0.738 | 0.141 | 0.492 |
| **01119** | 3+14 | 15.3 | 0.927 | 0.39 | 0.547 |
| **01120** | 3+16 | 18.8 | 1.693 | 0.766 | 0.48 |
| **01121** | 3+15 | 10.6 | 0.1608 | 0.816 | 1.416 |
| **01122** | 3+14 | 18.5 | 1.233 | 0.634 | 0.669 |
| **01123** | 4+17 | 17.9 | 3.2 | 0.475 | 3.2 |
| **01124** | 3+15 | 5.5 | 1.694 | 0.356 | 0.64 |
| **01125** | 3+16 | 52.8 | 1.7 | 0.614 | 1.286 |
| **01126** | 3+15 | 17.6 | 9.1 | 0.7 | 4 |
| **01127** | 4+15 | 14 | 8.6 | 0.5 | 3 |
| **01128** | 3+15 | 22.6 | 14.2 | 0.7 | 3.4 |
| **01129** | 4+14 | 15.8 | 19.3 | 1.1 | 4.1 |
| **01130** | 3+15 | 16.4 | 9.5 | 0.7 | 3 |

| | | | | | |
|---|---|---|---|---|---|
| **01131** | 3+16 | 13.7 | 1.074 | 0.43 | 0.654 |
| **01132** | 3+15 | 14.9 | 10.2 | 0.5 | 2.8 |
| **01133** | 4+14 | 27.7 | 3 | 0.6 | 2.9 |
| **01134** | 4+15 | 30.9 | 19.6 | 1.5 | 5.4 |
| **01135** | 3+16 | 17 | 11.2 | 0.8 | 3.8 |
| **01136** | 5+17 | 20.7 | 3 | 0.4 | 3 |
| **01137** | 4+14 | 12.2 | 8.9 | 0.6 | 3.6 |
| **01138** | 4+15 | 35.4 | 56.5 | 1 | 6.6 |
| **01139** | 4+15 | 10.1 | 12.8 | 0.7 | 1.3 |
| **01140** | 4+15 | 9.9 | 2.8 | 0.467 | 1.9 |
| **01141** | 3+17 | 15.2 | 3.1 | 0.6 | 2.2 |
| **01142** | 3+16 | 19.7 | 23.1 | 0.8 | 4.8 |
| **01143** | 3+15 | 12.7 | 1.1 | 0.6 | 0.8 |
| **01476** | 3+16 | 27.1 | 7 | 1 | 2.4 |
| **01477** | 4+17 | 20.6 | 7.1 | 1 | 3.4 |
| **01478** | 3+16 | 20.9 | 5 | 2.1 | 4.1 |
| **01700** | 4+15 | 15.6 | 4.4 | 0.8 | 37.4 |
| **01701** | 3+15 | 22.1 | 6.1 | 1.8 | 6.8 |
| **01702** | 3+17 | 16.1 | 3.2 | 1.8 | 46.8 |
| **01703** | 3+16 | 13.8 | 2.2 | 0.8 | 58.3 |
| **01704** | 3+15 | 4.2 | 2.3 | 0.231 | 0.827 |
| **01705** | 3+15 | 39.1 | 2.4 | 2.9 | 60.1 |
| **01706** | 3+16 | 9.1 | 3.2 | 1.1 | 4.7 |
| **01707** | 3+14 | 18 | 4.3 | 1.2 | 68.1 |
| **01708** | 3+16 | 28.1 | 7.4 | 1.7 | 72.1 |
| **01709** | 3+16 | 14.7 | 3.5 | 0.7 | 30.5 |
| **01710** | 3+17 | 34.8 | 4.8 | 2 | 107.7 |
| **01711** | 3+18 | 23.5 | 11.2 | 3.5 | 71 |
| **01712** | 4+17 | 16.5 | 6.5 | 1.2 | 46.2 |
| **01713** | 3+18 | 20.7 | 6.9 | 3.2 | 87.8 |
| **01714** | 3+18 | 9.9 | 3.9 | 1 | 35.4 |
| **01715** | 3+15 | 25.8 | 4.5 | 1.5 | 1.9 |
| **01716** | 3+16 | 30.5 | 5 | 3.5 | 113.6 |
| **01717** | 3+16 | 34.4 | 3.6 | 2 | 51.7 |
| **01718** | 4+16 | 12.1 | 2.6 | 1 | 3.6 |
| **01719** | 3+15 | 5.1 | 0.619 | 0.167 | 0.653 |
| **01720** | 3+14 | 20.5 | 4.1 | 2 | 6.9 |
| **01721** | 4+17 | 9.2 | 14.5 | 1.7 | 16.2 |
| **01722** | 3+16 | 20.5 | 6.1 | 2.1 | 41.3 |
| **01723** | 4+17 | 28.3 | 3.8 | 1 | 69 |
| **01724** | 3+18 | 37.5 | 8.4 | 1.2 | 74.2 |
| **01725** | 3+16 | 26.3 | 11.9 | 1.4 | 85 |
| **01826** | 4+16 | 5 | 2.2 | 0.398 | 2.6 |
| **01827** | 4+17 | 14.5 | 1.5 | 0.37 | 1.165 |
| **01828** | 4+15 | 8.1 | 1.09 | 0.313 | 0.688 |
| **01829** | 4+15 | 13.3 | 12.5 | 0.56 | 3 |
| **01830** | 3+14 | 14.1 | 4 | 0.496 | 2.8 |
| **01831** | 3+15 | 10.7 | 5.3 | 0.4 | 2.2 |
| **01832** | 3+16 | 23.3 | 7.2 | 1.6 | 7.1 |
| **01833** | 4+15 | 8.1 | 7.2 | 0.6 | 2.7 |
| **01834** | 4+16 | 20.4 | 3.2 | 0.4 | 1.8 |
| **01835** | 4+16 | 7 | 14.7 | 0.553 | 3 |
| **01836** | 5+19 | 19.5 | 6.2 | 0.448 | 3.5 |
| **01837** | 3+17 | 16.4 | 11 | 0.7 | 21.9 |
| **01838** | 3+14 | 10.9 | 7.6 | 0.47 | 1.8 |

| | | | | | |
|---|---|---|---|---|---|
| **01839** | 4+17 | 10.8 | 11.5 | 0.48 | 2.4 |
| **01840** | 4+16 | 9.9 | 1.352 | 0.16 | 0.356 |
| **01841** | 3+16 | 20 | 14 | 0.8 | 3.4 |
| **01842** | 3+17 | 13.8 | 20 | 0.8 | 4.5 |
| **01843** | 3+14 | 33.4 | 22 | 1.3 | 8.7 |
| **01844** | 4+17 | 38.9 | 4.8 | 0.9 | 4.1 |
| **01845** | 3+18 | 8.9 | 6.9 | 0.618 | 3.5 |
| **01846** | 3+14 | 21.3 | 4.8 | 0.6 | 2.9 |
| **01847** | 4+17 | 17.5 | 7.1 | 0.52 | 3 |
| **01848** | 3+16 | 6.1 | 6.8 | 0.366 | 2.3 |
| **01849** | 3+16 | 11 | 11.4 | 0.5 | 2.9 |
| **01850** | 3+15 | 20.3 | 3 | 0.29 | 0.968 |

Appendix 3: Additional details for the 139 *C. rupestris* samples. Age validated by otolith sectioning, total fish mass habitat depth, and trawl coordinates are displayed for each individual, identified by unique numbers (ID No.).

| ID No. | Age | Fish mass (g) | Habitat depth (m) | Latitude | Longitude |
|---|---|---|---|---|---|
| **01026** | 17 | 375 | 1640 | 59.109 | -9.1358333 |
| **01027** | 16 | 170 | 1640 | 59.109 | -9.1358333 |
| **01028** | 20 | 525 | 1640 | 59.109 | -9.1358333 |
| **01029** | 23 | 580 | 1640 | 59.109 | -9.1358333 |
| **01030** | 24 | 395 | 1640 | 59.109 | -9.1358333 |
| **01031** | 19 | 360 | 1640 | 59.109 | -9.1358333 |
| **01032** | 30 | 610 | 1640 | 59.109 | -9.1358333 |
| **01033** | 20 | 515 | 1640 | 59.109 | -9.1358333 |
| **01034** | 19 | 320 | 1640 | 59.109 | -9.1358333 |
| **01035** | 23 | 420 | 1640 | 59.109 | -9.1358333 |
| **01036** | 16 | 240 | 1640 | 59.109 | -9.1358333 |
| **01037** | 16 | 175 | 1640 | 59.109 | -9.1358333 |
| **01038** | 31 | 610 | 1640 | 59.109 | -9.1358333 |
| **01039** | 34 | 545 | 1640 | 59.109 | -9.1358333 |
| **01040** | 32 | 1495 | 1400 | 59.3721667 | -8.6388333 |
| **01041** | 31 | 2280 | 1400 | 59.3721667 | -8.6388333 |
| **01042** | 23 | 895 | 1400 | 59.3721667 | -8.6388333 |
| **01043** | 25 | 430 | 1400 | 59.3721667 | -8.6388333 |
| **01044** | 30 | 940 | 1400 | 59.3721667 | -8.6388333 |
| **01045** | 43 | 1230 | 1400 | 59.3721667 | -8.6388333 |
| **01046** | 11 | 110 | 1400 | 59.3721667 | -8.6388333 |
| **01047** | | 395 | 1400 | 59.3721667 | -8.6388333 |
| **01048** | 29 | 735 | 1400 | 59.3721667 | -8.6388333 |
| **01049** | 29 | 1500 | 1400 | 59.3721667 | -8.6388333 |
| **01050** | 31 | 960 | 1400 | 59.3721667 | -8.6388333 |
| **01051** | 14 | 525 | 1400 | 59.3721667 | -8.6388333 |
| **01052** | 24 | 970 | 1400 | 59.3721667 | -8.6388333 |
| **01053** | 24 | 595 | 1400 | 59.3721667 | -8.6388333 |
| **01054** | 25 | 720 | 1400 | 59.3721667 | -8.6388333 |
| **01055** | 43 | 1735 | 1400 | 59.3721667 | -8.6388333 |
| **01056** | 26 | 935 | 1400 | 59.3721667 | -8.6388333 |
| **01057** | 28 | 425 | 1400 | 59.3721667 | -8.6388333 |
| **01058** | 20 | 615 | 1400 | 59.3721667 | -8.6388333 |
| **01059** | 29 | 720 | 1400 | 59.3721667 | -8.6388333 |
| **01060** | 15 | 110 | 1400 | 59.3721667 | -8.6388333 |
| **01061** | 15 | 175 | 1400 | 59.3721667 | -8.6388333 |
| **01062** | 22 | 185 | 1400 | 59.3721667 | -8.6388333 |

| | | | | | |
|---|---|---|---|---|---|
| **01063** | 17 | 260 | 1400 | 59.3721667 | -8.6388333 |
| **01064** | 12 | 80 | 1400 | 59.3721667 | -8.6388333 |
| **01065** | 16 | 30.5 | 1400 | 59.3721667 | -8.6388333 |
| **01066** | 22 | 34 | 1400 | 59.3721667 | -8.6388333 |
| **01067** | 6 | 28 | 1400 | 59.3721667 | -8.6388333 |
| **01068** | 13 | 30 | 1400 | 59.3721667 | -8.6388333 |
| **01101** | | 1960 | 1060 | 59.1058333 | -9.8571667 |
| **01102** | 32 | 1460 | 1060 | 59.1058333 | -9.8571667 |
| **01103** | 26 | 1350 | 1060 | 59.1058333 | -9.8571667 |
| **01105** | 31 | 1730 | 1060 | 59.1058333 | -9.8571667 |
| **01106** | 27 | 1310 | 1060 | 59.1058333 | -9.8571667 |
| **01107** | 28 | 1580 | 1060 | 59.1058333 | -9.8571667 |
| **01108** | 44 | 1060 | 1060 | 59.1058333 | -9.8571667 |
| **01109** | 29 | 1950 | 1060 | 59.1058333 | -9.8571667 |
| **01110** | 27 | 1405 | 1060 | 59.1058333 | -9.8571667 |
| **01111** | 27 | 1000 | 1060 | 59.1058333 | -9.8571667 |
| **01112** | 23 | 1020 | 1060 | 59.1058333 | -9.8571667 |
| **01113** | 7 | 14.8 | 1060 | 59.1058333 | -9.8571667 |
| **01114** | | 79.5 | 1060 | 59.1058333 | -9.8571667 |
| **01115** | 12 | 27 | 1060 | 59.1058333 | -9.8571667 |
| **01116** | 10 | 41.4 | 1060 | 59.1058333 | -9.8571667 |
| **01117** | 12 | 148.6 | 1060 | 59.1058333 | -9.8571667 |
| **01118** | 24 | 159.5 | 1060 | 59.1058333 | -9.8571667 |
| **01119** | 20 | 282.5 | 1060 | 59.1058333 | -9.8571667 |
| **01120** | 15 | 351 | 1060 | 59.1058333 | -9.8571667 |
| **01121** | 20 | 380.4 | 1060 | 59.1058333 | -9.8571667 |
| **01122** | 18 | 361.8 | 1060 | 59.1058333 | -9.8571667 |
| **01123** | 29 | 727 | 1060 | 59.1058333 | -9.8571667 |
| **01124** | 24 | 435.6 | 1060 | 59.1058333 | -9.8571667 |
| **01125** | 18 | 619 | 1060 | 59.1058333 | -9.8571667 |
| **01126** | 32 | 680 | 1830 | 58.6913333 | -9.9103333 |
| **01127** | 38 | 410 | 1830 | 58.6913333 | -9.9103333 |
| **01128** | 26 | 625 | 1830 | 58.6913333 | -9.9103333 |
| **01129** | 34 | 620 | 1830 | 58.6913333 | -9.9103333 |
| **01130** | 28 | 550 | 1830 | 58.6913333 | -9.9103333 |
| **01131** | 22 | 355 | 1830 | 58.6913333 | -9.9103333 |
| **01132** | 36 | 465 | 1830 | 58.6913333 | -9.9103333 |
| **01133** | 27 | 590 | 1830 | 58.6913333 | -9.9103333 |
| **01134** | | 890 | 1830 | 58.6913333 | -9.9103333 |
| **01135** | 31 | 610 | 1830 | 58.6913333 | -9.9103333 |
| **01136** | | 450 | 1830 | 58.6913333 | -9.9103333 |
| **01137** | 27 | 600 | 1830 | 58.6913333 | -9.9103333 |
| **01138** | 30 | 935 | 1830 | 58.6913333 | -9.9103333 |
| **01139** | | 680 | 1830 | 58.6913333 | -9.9103333 |
| **01140** | 27 | 375 | 1830 | 58.6913333 | -9.9103333 |
| **01141** | 26 | 395 | 1830 | 58.6913333 | -9.9103333 |
| **01142** | 45 | 670 | 1830 | 58.6913333 | -9.9103333 |
| **01143** | 21 | 335 | 1830 | 58.6913333 | -9.9103333 |
| **01476** | 27 | 1310 | 720 | 59.2091667 | -9.893 |
| **01477** | 25 | 1290 | 720 | 59.2091667 | -9.893 |
| **01478** | 22 | 1250 | 720 | 59.2091667 | -9.893 |
| **01700** | 20 | 1345 | 720 | 59.2091667 | -9.893 |
| **01701** | 26 | 1540 | 720 | 59.2091667 | -9.893 |
| **01702** | 23 | 1135 | 720 | 59.2091667 | -9.893 |
| **01703** | 29 | 1700 | 720 | 59.2091667 | -9.893 |

| | | | | | |
|---|---|---|---|---|---|
| **01704** | 22 | 295 | 720 | 59.2091667 | -9.893 |
| **01705** | 25 | 1490 | 720 | 59.2091667 | -9.893 |
| **01706** | 20 | 815 | 720 | 59.2091667 | -9.893 |
| **01707** | 24 | 1710 | 720 | 59.2091667 | -9.893 |
| **01708** | 32 | 1935 | 720 | 59.2091667 | -9.893 |
| **01709** | 26 | 1060 | 720 | 59.2091667 | -9.893 |
| **01710** | 31 | 1880 | 720 | 59.2091667 | -9.893 |
| **01711** | 28 | 2170 | 720 | 59.2091667 | -9.893 |
| **01712** | 31 | 1560 | 720 | 59.2091667 | -9.893 |
| **01713** | 23 | 1750 | 720 | 59.2091667 | -9.893 |
| **01714** | 24 | 1500 | 720 | 59.2091667 | -9.893 |
| **01715** | 23 | 1075 | 720 | 59.2091667 | -9.893 |
| **01716** | 27 | 1775 | 720 | 59.2091667 | -9.893 |
| **01717** | 21 | 1520 | 720 | 59.2091667 | -9.893 |
| **01718** | 21 | 680 | 720 | 59.2091667 | -9.893 |
| **01719** | 16 | 200 | 720 | 59.2091667 | -9.893 |
| **01720** | 31 | 1870 | 1060 | 59.1058333 | -9.8571667 |
| **01721** | 39 | 2410 | 1060 | 59.1058333 | -9.8571667 |
| **01722** | 30 | 1465 | 1060 | 59.1058333 | -9.8571667 |
| **01723** | 30 | 1430 | 1060 | 59.1058333 | -9.8571667 |
| **01724** | 32 | 1245 | 1060 | 59.1058333 | -9.8571667 |
| **01725** | 31 | 1680 | 1060 | 59.1058333 | -9.8571667 |
| **01826** | 19 | 330 | 1830 | 58.6913333 | -9.9103333 |
| **01827** | 22 | 320 | 1830 | 58.6913333 | -9.9103333 |
| **01828** | 23 | 280 | 1830 | 58.6913333 | -9.9103333 |
| **01829** | 33 | 425 | 1830 | 58.6913333 | -9.9103333 |
| **01830** | 24 | 405 | 1830 | 58.6913333 | -9.9103333 |
| **01831** | 28 | 420 | 1830 | 58.6913333 | -9.9103333 |
| **01832** | 29 | 840 | 1830 | 58.6913333 | -9.9103333 |
| **01833** | 33 | 555 | 1830 | 58.6913333 | -9.9103333 |
| **01834** | 23 | 350 | 1830 | 58.6913333 | -9.9103333 |
| **01835** | 39 | 650 | 1640 | 59.109 | -9.1358333 |
| **01836** | 30 | 410 | 1640 | 59.109 | -9.1358333 |
| **01837** | 37 | 720 | 1640 | 59.109 | -9.1358333 |
| **01838** | 22 | 405 | 1640 | 59.109 | -9.1358333 |
| **01839** | 31 | 560 | 1640 | 59.109 | -9.1358333 |
| **01840** | 16 | 225 | 1640 | 59.109 | -9.1358333 |
| **01841** | 27 | 660 | 1640 | 59.109 | -9.1358333 |
| **01842** | 36 | 820 | 1640 | 59.109 | -9.1358333 |
| **01843** | 32 | 935 | 1640 | 59.109 | -9.1358333 |
| **01844** | 28 | 820 | 1640 | 59.109 | -9.1358333 |
| **01845** | 26 | 440 | 1640 | 59.109 | -9.1358333 |
| **01846** | 23 | 475 | 1640 | 59.109 | -9.1358333 |
| **01847** | 24 | 600 | 1640 | 59.109 | -9.1358333 |
| **01848** | 27 | 485 | 1640 | 59.109 | -9.1358333 |
| **01849** | 30 | 555 | 1640 | 59.109 | -9.1358333 |
| **01850** | 26 | 385 | 1640 | 59.109 | -9.1358333 |

Appendix 4: Post-hoc Games-Howell test output for differences in pre-anal fin length across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | 10.022222 | 8.543401 | .7664082 | -14.33912 | 34.38356 |
| | 1400 | 29.750000 | 8.371422 | .0086607 | 5.77745 | 53.72255 |
| | 1640 | 47.851190 | 5.834421 | .0000000 | 31.04778 | 64.65460 |
| | 1830 | 43.885802 | 5.954020 | .0000001 | 26.78402 | 60.98759 |
| **1060** | 720 | -10.022222 | 8.543401 | .7664082 | -34.38356 | 14.33912 |
| | 1400 | 19.727778 | 9.565311 | .2549474 | -7.50528 | 46.96083 |
| | 1640 | 37.828968 | 7.446808 | .0001577 | 16.26934 | 59.38859 |
| | 1830 | 33.863580 | 7.540878 | .0007678 | 12.08237 | 55.64479 |
| **1400** | 720 | -29.750000 | 8.371422 | .0086607 | -53.72255 | -5.77745 |
| | 1060 | -19.727778 | 9.565311 | .2549474 | -46.96083 | 7.50528 |
| | 1640 | 18.101190 | 7.248858 | .1206265 | -3.04692 | 39.24930 |
| | 1830 | 14.135802 | 7.345464 | .3278517 | -7.23385 | 35.50546 |
| **1640** | 720 | -47.851190 | 5.834421 | .0000000 | -64.65460 | -31.04778 |
| | 1060 | -37.828968 | 7.446808 | .0001577 | -59.38859 | -16.26934 |
| | 1400 | -18.101190 | 7.248858 | .1206265 | -39.24930 | 3.04692 |
| | 1830 | -3.965388 | 4.232682 | .8812025 | -15.92202 | 7.99125 |
| **1830** | 720 | -43.885802 | 5.954020 | .0000001 | -60.98759 | -26.78402 |
| | 1060 | -33.863580 | 7.540878 | .0007678 | -55.64479 | -12.08237 |
| | 1400 | -14.135802 | 7.345464 | .3278517 | -35.50546 | 7.23385 |
| | 1640 | 3.965388 | 4.232682 | .8812025 | -7.99125 | 15.92202 |

Appendix 5: Post-hoc Games-Howell test output for differences in total mass across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | 187.556 | 153.431 | .7385099 | -249.62 | 624.73 |
| | 1400 | 516.223 | 147.566 | .0098314 | 94.33 | 938.12 |
| | 1640 | 876.461 | 99.103 | .0000000 | 586.25 | 1166.67 |
| | 1830 | 874.874 | 99.873 | .0000000 | 582.96 | 1166.78 |
| **1060** | 720 | -187.556 | 153.431 | .7385099 | -624.73 | 249.62 |
| | 1400 | 328.667 | 166.618 | .2964317 | -145.67 | 803.01 |
| | 1640 | 688.905 | 125.726 | .0000843 | 320.98 | 1056.83 |
| | 1830 | 687.318 | 126.334 | .0000884 | 318.07 | 1056.56 |
| **1400** | 720 | -516.223 | 147.566 | .0098314 | -938.12 | -94.33 |
| | 1060 | -328.667 | 166.618 | .2964317 | -803.01 | 145.67 |
| | 1640 | 360.238 | 118.498 | .0415374 | 10.24 | 710.24 |
| | 1830 | 358.651 | 119.142 | .0438185 | 7.31 | 710.00 |
| **1640** | 720 | -876.461 | 99.103 | .0000000 | -1166.67 | -586.25 |
| | 1060 | -688.905 | 125.726 | .0000843 | -1056.83 | -320.98 |
| | 1400 | -360.238 | 118.498 | .0415374 | -710.24 | -10.24 |
| | 1830 | -1.587 | 47.335 | .9999997 | -135.29 | 132.12 |
| **1830** | 720 | -874.874 | 99.873 | .0000000 | -1166.78 | -582.96 |
| | 1060 | -687.318 | 126.334 | .0000884 | -1056.56 | -318.07 |
| | 1400 | -358.651 | 119.142 | .0438185 | -710.00 | -7.31 |
| | 1640 | 1.587 | 47.335 | .9999997 | -132.12 | 135.29 |

Appendix 6: Post-hoc Games-Howell test output for differences in relative liver mass across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | -.0085031154 | .0041134184 | .2636572 | -.020536384 | .003530153 |
| | 1400 | -.0150426114 | .0031439684 | .0005586 | -.024267844 | -.005817379 |
| | 1640 | -.0202747367 | .0033849410 | .0000095 | -.030038855 | -.010510619 |
| | 1830 | -.0169211165 | .0022421063 | .0000000 | -.023327324 | -.010514909 |
| **1060** | 720 | .0085031154 | .0041134184 | .2636572 | -.003530153 | .020536384 |
| | 1400 | -.0065394959 | .0049460064 | .6792231 | -.020649139 | .007570147 |
| | 1640 | -.0117716213 | .0051025746 | .1611344 | -.026256088 | .002712845 |
| | 1830 | -.0084180011 | .0044278079 | .3362087 | -.021171236 | .004335234 |
| **1400** | 720 | .0150426114 | .0031439684 | .0005586 | .005817379 | .024267844 |
| | 1060 | .0065394959 | .0049460064 | .6792231 | -.007570147 | .020649139 |
| | 1640 | -.0052321254 | .0043589671 | .7511538 | -.017598458 | .007134207 |
| | 1830 | -.0018785051 | .0035453926 | .9836718 | -.012054269 | .008297259 |
| **1640** | 720 | .0202747367 | .0033849410 | .0000095 | .010510619 | .030038855 |
| | 1060 | .0117716213 | .0051025746 | .1611344 | -.002712845 | .026256088 |
| | 1400 | .0052321254 | .0043589671 | .7511538 | -.007134207 | .017598458 |
| | 1830 | .0033536202 | .0037607309 | .8984540 | -.007336870 | .014044111 |
| **1830** | 720 | .0169211165 | .0022421063 | .0000000 | .010514909 | .023327324 |
| | 1060 | .0084180011 | .0044278079 | .3362087 | -.004335234 | .021171236 |
| | 1400 | .0018785051 | .0035453926 | .9836718 | -.008297259 | .012054269 |
| | 1640 | -.0033536202 | .0037607309 | .8984540 | -.014044111 | .007336870 |

Appendix 7: Post-hoc Games-Howell test output for differences in relative swimbladder mass across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | -.0018082654 | .0007622707 | .1496770 | -.004013055 | .000396525 |
| | 1400 | -.0037105732 | .0012116705 | .0400835 | -.007295839 | -.000125307 |
| | 1640 | -.0086563102 | .0013008085 | .0000021 | -.012428431 | -.004884190 |
| | 1830 | -.0121640748 | .0024318366 | .0002768 | -.019269681 | -.005058469 |
| **1060** | 720 | .0018082654 | .0007622707 | .1496770 | -.000396525 | .004013055 |
| | 1400 | -.0019023078 | .0013643620 | .6354958 | -.005837031 | .002032416 |
| | 1640 | -.0068480448 | .0014441056 | .0002332 | -.010965877 | -.002730213 |
| | 1830 | -.0103558094 | .0025114075 | .0023154 | -.017636018 | -.003075601 |
| **1400** | 720 | .0037105732 | .0012116705 | .0400835 | .000125307 | .007295839 |
| | 1060 | .0019023078 | .0013643620 | .6354958 | -.002032416 | .005837031 |
| | 1640 | -.0049457370 | .0017241026 | .0463224 | -.009837198 | -.000054275 |
| | 1830 | -.0084535016 | .0026822111 | .0250645 | -.016142513 | -.000764490 |
| **1640** | 720 | .0086563102 | .0013008085 | .0000021 | .004884190 | .012428431 |
| | 1060 | .0068480448 | .0014441056 | .0002332 | .002730213 | .010965877 |
| | 1400 | .0049457370 | .0017241026 | .0463224 | .000054275 | .009837198 |
| | 1830 | -.0035077646 | .0027236398 | .7000164 | -.011292393 | .004276864 |
| **1830** | 720 | .0121640748 | .0024318366 | .0002768 | .005058469 | .019269681 |
| | 1060 | .0103558094 | .0025114075 | .0023154 | .003075601 | .017636018 |
| | 1400 | .0084535016 | .0026822111 | .0250645 | .000764490 | .016142513 |
| | 1640 | .0035077646 | .0027236398 | .7000164 | -.004276864 | .011292393 |

Appendix 8: Post-hoc Tukey test output for differences in relative maximum body width across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | .0201166742 | .0085793924 | .1383527 | -.003654927 | .043888276 |
| | 1400 | .0253357961 | .0088676352 | .0397749 | .000765538 | .049906054 |
| | 1640 | .0376246153 | .0082811228 | .0001306 | .014679453 | .060569777 |
| | 1830 | .0363074985 | .0083483259 | .0002802 | .013176131 | .059438866 |
| **1060** | 720 | -.0201166742 | .0085793924 | .1383527 | -.043888276 | .003654927 |
| | 1400 | .0052191219 | .0086853149 | .9746916 | -.018845967 | .029284211 |
| | 1640 | .0175079411 | .0080855881 | .2002725 | -.004895438 | .039911320 |
| | 1830 | .0161908242 | .0081544029 | .2795000 | -.006403225 | .038784873 |
| **1400** | 720 | -.0253357961 | .0088676352 | .0397749 | -.049906054 | -.000765538 |
| | 1060 | -.0052191219 | .0086853149 | .9746916 | -.029284211 | .018845967 |
| | 1640 | .0122888192 | .0083908113 | .5874322 | -.010960266 | .035537904 |
| | 1830 | .0109717024 | .0084571429 | .6933768 | -.012461172 | .034404577 |
| **1640** | 720 | -.0376246153 | .0082811228 | .0001306 | -.060569777 | -.014679453 |
| | 1060 | -.0175079411 | .0080855881 | .2002725 | -.039911320 | .004895438 |
| | 1400 | -.0122888192 | .0083908113 | .5874322 | -.035537904 | .010960266 |
| | 1830 | -.0013171169 | .0078399813 | .9998189 | -.023039973 | .020405740 |
| **1830** | 720 | -.0363074985 | .0083483259 | .0002802 | -.059438866 | -.013176131 |
| | 1060 | -.0161908242 | .0081544029 | .2795000 | -.038784873 | .006403225 |
| | 1400 | -.0109717024 | .0084571429 | .6933768 | -.034404577 | .012461172 |
| | 1640 | .0013171169 | .0078399813 | .9998189 | -.020405740 | .023039973 |

Appendix 9: Post-hoc Tukey test output for differences in relative head length across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | .0034806105 | .0086443351 | .9943960 | -.020470933 | .027432154 |
| | 1400 | -.0338480037 | .0089347597 | .0022074 | -.058604249 | -.009091758 |
| | 1640 | -.0210458527 | .0083438077 | .0926422 | -.044164701 | .002072995 |
| | 1830 | -.0138708627 | .0084115195 | .4694942 | -.037177325 | .009435600 |
| **1060** | 720 | -.0034806105 | .0086443351 | .9943960 | -.027432154 | .020470933 |
| | 1400 | -.0373286142 | .0087510593 | .0003861 | -.061575867 | -.013081362 |
| | 1640 | -.0245264632 | .0081467929 | .0260209 | -.047099427 | -.001953500 |
| | 1830 | -.0173514732 | .0082161285 | .2220122 | -.040116550 | .005413604 |
| **1400** | 720 | .0338480037 | .0089347597 | .0022074 | .009091758 | .058604249 |
| | 1060 | .0373286142 | .0087510593 | .0003861 | .013081362 | .061575867 |
| | 1640 | .0128021510 | .0084543265 | .5553627 | -.010622920 | .036227222 |
| | 1830 | .0199771410 | .0085211602 | .1384578 | -.003633112 | .043587394 |
| **1640** | 720 | .0210458527 | .0083438077 | .0926422 | -.002072995 | .044164701 |
| | 1060 | .0245264632 | .0081467929 | .0260209 | .001953500 | .047099427 |
| | 1400 | -.0128021510 | .0084543265 | .5553627 | -.036227222 | .010622920 |
| | 1830 | .0071749900 | .0078993269 | .8932067 | -.014712300 | .029062280 |
| **1830** | 720 | .0138708627 | .0084115195 | .4694942 | -.009435600 | .037177325 |
| | 1060 | .0173514732 | .0082161285 | .2220122 | -.005413604 | .040116550 |
| | 1400 | -.0199771410 | .0085211602 | .1384578 | -.043587394 | .003633112 |
| | 1640 | -.0071749900 | .0078993269 | .8932067 | -.029062280 | .014712300 |

Appendix 10: Post-hoc Games-Howell test output for differences in relative orbit length across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | -.0097544955 | .0059878056 | .4882049 | -.026851594 | .007342603 |
| | 1400 | -.0116312445 | .0050469305 | .1643733 | -.026037480 | .002774991 |
| | 1640 | -.0229251461 | .0047946310 | .0001642 | -.036523923 | -.009326369 |
| | 1830 | -.0161983502 | .0041151286 | .0029644 | -.027978616 | -.004418084 |
| **1060** | 720 | .0097544955 | .0059878056 | .4882049 | -.007342603 | .026851594 |
| | 1400 | -.0018767490 | .0061690967 | .9980591 | -.019470802 | .015717304 |
| | 1640 | -.0131706507 | .0059644558 | .1967526 | -.030171850 | .003830549 |
| | 1830 | -.0064438547 | .0054332799 | .7591686 | -.022121080 | .009233371 |
| **1400** | 720 | .0116312445 | .0050469305 | .1643733 | -.002774991 | .026037480 |
| | 1060 | .0018767490 | .0061690967 | .9980591 | -.015717304 | .019470802 |
| | 1640 | -.0112939017 | .0050192056 | .1807542 | -.025565939 | .002978135 |
| | 1830 | -.0045671057 | .0043747253 | .8331089 | -.017156806 | .008022595 |
| **1640** | 720 | .0229251461 | .0047946310 | .0001642 | .009326369 | .036523923 |
| | 1060 | .0131706507 | .0059644558 | .1967526 | -.003830549 | .030171850 |
| | 1400 | .0112939017 | .0050192056 | .1807542 | -.002978135 | .025565939 |
| | 1830 | .0067267960 | .0040810783 | .4750166 | -.004845565 | .018299157 |
| **1830** | 720 | .0161983502 | .0041151286 | .0029644 | .004418084 | .027978616 |
| | 1060 | .0064438547 | .0054332799 | .7591686 | -.009233371 | .022121080 |
| | 1400 | .0045671057 | .0043747253 | .8331089 | -.008022595 | .017156806 |
| | 1640 | -.0067267960 | .0040810783 | .4750166 | -.018299157 | .004845565 |

Appendix 11: Post-hoc Tukey test output for differences in relative inter-orbital width across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | .0077350742 | .0047918270 | .4913950 | -.005542020 | .021012168 |
| | 1400 | -.0147071641 | .0049528186 | .0292173 | -.028430330 | -.000983998 |
| | 1640 | -.0021426751 | .0046252352 | .9904222 | -.014958180 | .010672830 |
| | 1830 | -.0028337910 | .0046627700 | .9736225 | -.015753297 | .010085715 |
| **1060** | 720 | -.0077350742 | .0047918270 | .4913950 | -.021012168 | .005542020 |
| | 1400 | -.0224422383 | .0048509876 | .0000936 | -.035883253 | -.009001223 |
| | 1640 | -.0098777493 | .0045160237 | .1917934 | -.022390653 | .002635155 |
| | 1830 | -.0105688652 | .0045544586 | .1457407 | -.023188264 | .002050534 |
| **1400** | 720 | .0147071641 | .0049528186 | .0292173 | .000983998 | .028430330 |
| | 1060 | .0224422383 | .0048509876 | .0000936 | .009001223 | .035883253 |
| | 1640 | .0125644890 | .0046864993 | .0629378 | -.000420765 | .025549743 |
| | 1830 | .0118733731 | .0047235473 | .0945499 | -.001214533 | .024961279 |
| **1640** | 720 | .0021426751 | .0046252352 | .9904222 | -.010672830 | .014958180 |
| | 1060 | .0098777493 | .0045160237 | .1917934 | -.002635155 | .022390653 |
| | 1400 | -.0125644890 | .0046864993 | .0629378 | -.025549743 | .000420765 |
| | 1830 | -.0006911159 | .0043788455 | .9998587 | -.012823930 | .011441698 |
| **1830** | 720 | .0028337910 | .0046627700 | .9736225 | -.010085715 | .015753297 |
| | 1060 | .0105688652 | .0045544586 | .1457407 | -.002050534 | .023188264 |
| | 1400 | -.0118733731 | .0047235473 | .0945499 | -.024961279 | .001214533 |
| | 1640 | .0006911159 | .0043788455 | .9998587 | -.011441698 | .012823930 |

Appendix 12: Post-hoc Games-Howell test output for differences in relative gape size across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | -.0079960865 | .0065860692 | .7432991 | -.026790243 | .010798070 |
| | 1400 | .0340826941 | .0171019440 | .3016460 | -.016624955 | .084790343 |
| | 1640 | -.0349352612 | .0070164276 | .0001015 | -.054907714 | -.014962809 |
| | 1830 | -.0378737864 | .0061821602 | .0000019 | -.055428898 | -.020318675 |
| **1060** | 720 | .0079960865 | .0065860692 | .7432991 | -.010798070 | .026790243 |
| | 1400 | .0420787806 | .0174826480 | .1476884 | -.009414376 | .093571937 |
| | 1640 | -.0269391747 | .0078991612 | .0109445 | -.049310033 | -.004568317 |
| | 1830 | -.0298776999 | .0071683748 | .0011641 | -.050198182 | -.009557218 |
| **1400** | 720 | -.0340826941 | .0171019440 | .3016460 | -.084790343 | .016624955 |
| | 1060 | -.0420787806 | .0174826480 | .1476884 | -.093571937 | .009414376 |
| | 1640 | -.0690179553 | .0176492756 | .0051611 | -.120861908 | -.017174002 |
| | 1830 | -.0719564805 | .0173345256 | .0031441 | -.123135786 | -.020777175 |
| **1640** | 720 | .0349352612 | .0070164276 | .0001015 | .014962809 | .054907714 |
| | 1060 | .0269391747 | .0078991612 | .0109445 | .004568317 | .049310033 |
| | 1400 | .0690179553 | .0176492756 | .0051611 | .017174002 | .120861908 |
| | 1830 | -.0029385252 | .0075656821 | .9950280 | -.024345671 | .018468621 |
| **1830** | 720 | .0378737864 | .0061821602 | .0000019 | .020318675 | .055428898 |
| | 1060 | .0298776999 | .0071683748 | .0011641 | .009557218 | .050198182 |
| | 1400 | .0719564805 | .0173345256 | .0031441 | .020777175 | .123135786 |
| | 1640 | .0029385252 | .0075656821 | .9950280 | -.018468621 | .024345671 |

Appendix 13: Post-hoc Tukey test output for differences in relative gape length across five depth groups. Note that the critical value is *p*=0.003125.

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | .0030926380 | .0053188346 | .9775719 | -.011646670 | .017831946 |
| | 1400 | .0162311327 | .0054932688 | .0304824 | .001008441 | .031453825 |
| | 1640 | -.0145020114 | .0051384824 | .0437104 | -.028741536 | -.000262487 |
| | 1830 | .0054161871 | .0051791034 | .8334365 | -.008935904 | .019768279 |
| **1060** | 720 | -.0030926380 | .0053188346 | .9775719 | -.017831946 | .011646670 |
| | 1400 | .0131384946 | .0053188346 | .1046521 | -.001600813 | .027877803 |
| | 1640 | -.0175946494 | .0049515655 | .0049141 | -.031316199 | -.003873100 |
| | 1830 | .0023235491 | .0049937073 | .9902602 | -.011514781 | .016161879 |
| **1400** | 720 | -.0162311327 | .0054932688 | .0304824 | -.031453825 | -.001008441 |
| | 1060 | -.0131384946 | .0053188346 | .1046521 | -.027877803 | .001600813 |
| | 1640 | -.0307331440 | .0051384824 | .0000002 | -.044972669 | -.016493620 |
| | 1830 | -.0108149455 | .0051791034 | .2322010 | -.025167037 | .003537146 |
| **1640** | 720 | .0145020114 | .0051384824 | .0437104 | .000262487 | .028741536 |
| | 1060 | .0175946494 | .0049515655 | .0049141 | .003873100 | .031316199 |
| | 1400 | .0307331440 | .0051384824 | .0000002 | .016493620 | .044972669 |
| | 1830 | .0199181985 | .0048011574 | .0006036 | .006613453 | .033222944 |
| **1830** | 720 | -.0054161871 | .0051791034 | .8334365 | -.019768279 | .008935904 |
| | 1060 | -.0023235491 | .0049937073 | .9902602 | -.016161879 | .011514781 |
| | 1400 | .0108149455 | .0051791034 | .2322010 | -.003537146 | .025167037 |
| | 1640 | -.0199181985 | .0048011574 | .0006036 | -.033222944 | -.006613453 |

Appendix 14: Post-hoc Games-Howell test output for differences in the PCA extracted 'Size' component across five depth groups. Note that the critical value is $p$=0.01ȯ

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | .38709394 | .28213644 | .6484716 | -.4166036 | 1.1907915 |
| | 1400 | .96596185 | .28751396 | .0144570 | .1429821 | 1.7889416 |
| | 1640 | 1.56082229 | .20506841 | .0000001 | .9679346 | 2.1537100 |
| | 1830 | 1.47391742 | .20680464 | .0000004 | .8768354 | 2.0709994 |
| **1060** | 720 | -.38709394 | .28213644 | .6484716 | -1.1907915 | .4166036 |
| | 1400 | .57886791 | .31248088 | .3580936 | -.3111232 | 1.4688590 |
| | 1640 | 1.17372835 | .23881599 | .0002291 | .4839958 | 1.8634609 |
| | 1830 | 1.08682348 | .24030852 | .0006765 | .3935492 | 1.7800977 |
| **1400** | 720 | -.96596185 | .28751396 | .0144570 | -1.7889416 | -.1429821 |
| | 1060 | -.57886791 | .31248088 | .3580936 | -1.4688590 | .3111232 |
| | 1640 | .59486044 | .24514564 | .1380881 | -.1199006 | 1.3096215 |
| | 1830 | .50795557 | .24659986 | .2650362 | -.2101164 | 1.2260276 |
| **1640** | 720 | -1.56082229 | .20506841 | .0000001 | -2.1537100 | -.9679346 |
| | 1060 | -1.17372835 | .23881599 | .0002291 | -1.8634609 | -.4839958 |
| | 1400 | -.59486044 | .24514564 | .1380881 | -1.3096215 | .1199006 |
| | 1830 | -.08690487 | .14212763 | .9726503 | -.4885458 | .3147361 |
| **1830** | 720 | -1.47391742 | .20680464 | .0000004 | -2.0709994 | -.8768354 |
| | 1060 | -1.08682348 | .24030852 | .0006765 | -1.7800977 | -.3935492 |
| | 1400 | -.50795557 | .24659986 | .2650362 | -1.2260276 | .2101164 |
| | 1640 | .08690487 | .14212763 | .9726503 | -.3147361 | .4885458 |

Appendix 15: Post-hoc Games-Howell test output for differences in the PCA extracted 'Lipid Storage' component across five depth groups. Note that the critical value is $p$=0.01ȯ

| (I) Depth of Capture (m) | (J) Depth of Capture (m) | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| **720** | 1060 | -.41956390 | .18160390 | .1619509 | -.9372350 | .0981072 |
| | 1400 | -1.45408420 | .34526939 | .0024967 | -2.4692647 | -.4389037 |
| | 1640 | -.89835523 | .15439602 | .0000058 | -1.3372454 | -.4594650 |
| | 1830 | -.97327493 | .21236831 | .0003791 | -1.5786289 | -.3679209 |
| **1060** | 720 | .41956390 | .18160390 | .1619509 | -.0981072 | .9372350 |
| | 1400 | -1.03452030 | .35665100 | .0518139 | -2.0747937 | .0057531 |
| | 1640 | -.47879133 | .17840155 | .0728603 | -.9863419 | .0287593 |
| | 1830 | -.55371103 | .23041112 | .1321402 | -1.2069113 | .0994893 |
| **1400** | 720 | 1.45408420 | .34526939 | .0024967 | .4389037 | 2.4692647 |
| | 1060 | 1.03452030 | .35665100 | .0518139 | -.0057531 | 2.0747937 |
| | 1640 | .55572897 | .34359582 | .5010347 | -.4556488 | 1.5671067 |
| | 1830 | .48080928 | .37325629 | .7001793 | -.5980607 | 1.5596792 |
| **1640** | 720 | .89835523 | .15439602 | .0000058 | .4594650 | 1.3372454 |
| | 1060 | .47879133 | .17840155 | .0728603 | -.0287593 | .9863419 |
| | 1400 | -.55572897 | .34359582 | .5010347 | -1.5671067 | .4556488 |
| | 1830 | -.07491969 | .20963644 | .9963718 | -.6722570 | .5224176 |
| **1830** | 720 | .97327493 | .21236831 | .0003791 | .3679209 | 1.5786289 |
| | 1060 | .55371103 | .23041112 | .1321402 | -.0994893 | 1.2069113 |
| | 1400 | -.48080928 | .37325629 | .7001793 | -1.5596792 | .5980607 |
| | 1640 | .07491969 | .20963644 | .9963718 | -.5224176 | .6722570 |

Appendix 16: Table illustrating the positioning of non-synonymous SNPs across the OBSL1 coding region in *C. rupestris.*

| | SNP Position on Contig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Depth (m) | 286684 | 287258 | 293186 | 294904 | 296918 | 297048 | 297142 | 297969 | 298007 | 299069 | 299438 | 299556 | 299592 | 300971 | 301064 | 303529 | 304579 | 304819 | 305601 | 306399 | 306400 | 317058 | 317820 | 319011 | 320472 | 322143 | 322206 | 322635 | 322644 | 322995 | 323834 | 323846 |
| 750 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 750 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AG | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 750 | CA | CC | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 750 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | CT | AG | GG | AA | CA | TT | GG |
| 750 | CC | CC | GG | AT | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 750 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AG | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 750 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 750 | CA | CC | TG | AA | TT | TT | AA | AG | GC | AA | AG | TG | AT | GG | GG | TG | CT | GG | TT | AA | CC | AA | GG | AA | CA | CT | AG | GG | AA | CA | TT | GG |
| 750 | CC | CC | TG | AT | TG | CT | CA | AG | GC | AG | AG | TG | AT | GC | AA | TG | CT | AG | CT | AG | CT | AG | GC | AG | CA | CT | AG | AG | CA | CA | CT | AG |
| 750 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 750 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | CT | AG |
| 750 | CC | CC | TG | AT | TG | CT | CA | AG | GC | AG | AG | TG | AT | GC | AG | TG | CT | AG | CT | AG | CT | AG | GC | AG | CA | CT | AG | AG | CA | CA | CT | AG |
| 750 | CA | CC | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 750 | CA | CT | TG | AT | TG | CT | CA | AG | GC | GG | AG | TG | AT | GC | AA | TT | CC | AA | CC | GG | TT | AG | CC | AA | CA | CT | AG | AG | CA | CC | CT | AG |
| 1000 | CC | CC | GG | TT | GG | CC | CA | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1000 | AA | TT | TG | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1000 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1000 | CA | CC | TG | AT | TG | CT | CA | AA | GC | AG | AG | TG | AA | GC | AG | TG | CT | AG | TT | AG | CT | AA | GC | AA | CA | CT | AG | AG | CA | CA | CT | AG |
| 1000 | CC | CC | GG | AT | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | CA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | AG | CA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | CT | AA | CC | AA | GG | AA | AA | TT | AG | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AT | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AG | GG | AA | CA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AG | GC | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | CT | AG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AA | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | CT | AG | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AT | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AG | GG | AA | AA | TT | AG | GG | AA | CA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | CT | AA | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | CT | AA | GG | AA | AA | TT | GG |
| 1000 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | TT | AG | CT | AA | GC | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1500 | CC | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1500 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1500 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1500 | AA | CT | TG | TT | GG | CC | CA | GG | CC | GG | AG | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1500 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AA | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | CT | AG |
| 1500 | CC | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AG | CA | CC | CC | AA |
| 1500 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | GG | GG | TT | GG | N | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1500 | CA | CT | TG | AT | TG | CT | CA | AG | GC | AG | AG | TG | AT | GC | AG | TG | CT | AG | CT | AG | CT | AG | GC | AG | CA | CT | AG | AG | CA | CC | CT | AG |
| 1500 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | GC | AG | TT | CC | AA | CC | GG | TT | AG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1500 | CA | CT | TG | AT | TG | CT | CA | AG | GC | AG | AG | TG | AT | GC | AG | TG | CC | AA | CC | GG | TT | AG | CC | AA | CA | CT | AG | AG | CA | AA | CT | AG |
| 1500 | CA | CT | TG | AT | TG | CT | CA | AG | GC | AG | AA | TG | AT | GC | AA | TG | CT | AG | CT | AG | CT | AG | GC | AG | CA | CT | AG | GG | CA | CA | CT | AG |
| 1500 | CC | CC | GG | AA | TG | TT | AA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1500 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AA | GG | AG | GG | TT | GG | TT | AA | CC | AA | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1500 | CC | CC | GG | AA | TT | TT | AA | AA | GG | AA | GG | TT | AT | GG | AG | TG | CC | AA | CT | AA | CC | AG | GG | AA | AA | TT | AA | GG | AA | AA | TT | GG |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | N | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | N | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | AG | CT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CC | CC | TG | AT | TG | CT | CA | AG | GC | AG | AG | TG | AT | GC | AA | TG | CT | AG | CT | AG | CT | AG | GC | AA | AA | CT | AG | AG | CA | CA | CT | AG |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | CT | TG | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | CA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | TT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | GG | CC | CC | GG | AA | CC | CC | CC | AA |
| 1800 | AA | CT | TT | TT | GG | CC | CC | GG | CC | GG | AA | GG | TT | CC | AA | TT | CC | AA | CC | GG | TT | GG | CC | AG | CC | CC | GG | AA | CC | CC | CC | AA |

# Bibliography


Agarkova, I. & Perriard, J.C. 2005. The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere. *Trends in Cell Biology,* 15, 477-85.

Akbal-Delibas, B., Jagodzinski, F. & Haspel, N. 2013. A conservation and rigidity based method for detecting critical protein residues. *BMC Structural Biology,* 13 Suppl 1**,** S6.

Allain, V. 2001. Reproductive strategies of three deep-water benthopelagic fishes from the northeast Atlantic Ocean. *Fisheries Research,* 51**,** 165-176.

Angel, M. V. 1997. What is The Deep Sea? *In:* Randall, D. J. & Farrell, A. P. (eds.) *Deep-Sea Fishes.* Academic Press.

Atkinson, D. 1995. The biology and fishery of roundnose grenadier (*Coryphaenoides rupestris* Gunnerus, 1765) in the north west Atlantic. *In:* Hopper, A. (ed.) *Deep-water fisheries of the north Atlantic oceanic slope.* Springer.

Baco, A. R., Etter, R. J., Ribeiro, P. A., Von Der Heyden, S., Beerli, P. & Kinlan, B. P. 2016. A synthesis of genetic connectivity in deep-sea fauna and implications for marine reserve design. *Molecular Ecology,* 25**,** 3276-98.

Baker, K. D., Devine, J. A. & Haedrich, R. L. 2009. Deep-sea fishes in Canada's Atlantic: population declines and predicted recovery times. *Environmental biology of fishes,* 85**,** 79.

Beamish, R. 1979. Differences in the age of Pacific hake (*Merluccius productus*) using whole otoliths and sections of otoliths. *Journal of the Fisheries Board of Canada,* 36**,** 141-151.

Bergstad, O. 1990. Distribution, population structure, growth and reproduction of the roundnose grenadier *Coryphaenoides rupestris* (Pisces: Macrouridae) in the deep waters of the Skagerrak. *Marine Biology,* 107**,** 25-39.

Bergstad, O. A., Gjelsvik, G., Schander, C. & Hoines, A. S. 2010. Feeding ecology of *Coryphaenoides rupestris* from the mid-Atlantic Ridge. *PLoS One,* 5**,** e10453.

Bergstad, O. A., Øverbø Hansen, H. & Jørgensen, T. 2014. Intermittent recruitment and exploitation pulse underlying temporal variability in a demersal deep-water fish population. *ICES Journal of Marine Science,* 71**,** 2088-2100.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D. & Zardecki, C. 2002. The Protein Data Bank. *Acta Crystallographica Section D: Biological Crystallography,* 58**,** 899-907.

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R. & Bendayan, R. 2017. Non-normal data: Is ANOVA still a valid option? *Psicothema,* 29**,** 552-557.

Bogomolovas, J., Fleming, J. R., Anderson, B. R., Williams, R., Lange, S., Simon, B., Khan, M. M., Rudolf, R., Franke, B., Bullard, B., Rigden, D. J., Granzier, H., Labeit, S. & Mayans, O. 2016. Exploration of pathomechanisms triggered by a single-nucleotide polymorphism in titin's I-band: the cardiomyopathy-linked mutation T2580I. *Open Biology,* 6.

Bowering, W. R. & Nedreaas, K. H. 2000. A comparison of Greenland halibut (*Reinhardtius hippoglossoides* (Walbaum)) fisheries and distribution in the Northwest and Northeast Atlantic. *Sarsia,* 85**,** 61-76.

Brindley, A. A., Pickersgill, R. W., Partridge, J. C., Dunstan, D. J., Hunt, D. M. & Warren, M. J. 2008. Enzyme sequence and its relationship to hyperbaric stability of artificial and natural fish lactate dehydrogenases. *PLoS One,* 3**,** e2042.

Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., Goodstein, D. M., Elsik, C. G., Lewis, S. E., Stein, L. & Holmes, I. H. 2016. JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biology,* 17**,** 66.

Buesseler, K. O., Lamborg, C. H., Boyd, P. W., Lam, P. J., Trull, T. W., Bidigare, R. R., Bishop, J. K., Casciotti, K. L., Dehairs, F. & Elskens, M. 2007. Revisiting carbon flux through the ocean's twilight zone. *Science,* 316**,** 567-570.

Childress, J., Taylor, S., Cailliet, G. M. & Price, M. 1980. Patterns of growth, energy utilization and reproduction in some meso-and bathypelagic fishes off southern California. *Marine Biology,* 61**,** 27-40.

Childress, J. J. 1995. Are there physiological and biochemical adaptations of metabolism in deep-sea animals? *Trends in Ecology & Evolution,* 10**,** 30-6.

Cohen, D. M., Inada, T., Iwamoto, T. & Scialabba, N. 1990. Gadiform fishes of the world. *FAO Fisheries Synopsis,* 10**,** I.

Collins, A. & Anderson, T. 1995. The regulation of endogeneous energy stores during starvation and refeeding in the somatic tissues of the golden perch. *Journal of Fish Biology,* 47**,** 1004-1015.

Collins, M. A., Bailey, D. M., Ruxton, G. D. & Priede, I. G. 2005. Trends in body size across an environmental gradient: a differential response in scavenging and non-scavenging demersal deep-sea fish. *Proceedings of the Royal Society B: Biological Sciences,* 272**,** 2051-7.
Consortium, U. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research,* 47**,** D506-D515.
Cook, R., Fernandes, P., Florin, A., Lorance, P. & Nedreaas, K. 2015. *Coryphaenoides rupestris* [Online]. [Accessed 11 March 2020].
Culver, S. J. & Buzas, M. A. 2000. Global latitudinal species diversity gradient in deep-sea benthic foraminifera. *Deep Sea Research Part I: Oceanographic Research Papers,* 47**,** 259-275.
Danovaro, R., Snelgrove, P. V. & Tyler, P. 2014. Challenging the paradigms of deep-sea ecology. *Trends in Ecology & Evolution,* 29**,** 465-75.
Delano, W. L. 2002. The PyMOL molecular graphics system. *Schrodinger*.
Delong, E. F. & Yayanos, A. A. 1985. Adaptation of the membrane lipids of a deep-sea bacterium to changes in hydrostatic pressure. *Science,* 228**,** 1101-3.
Devine, J. A., Baker, K. D. & Haedrich, R. L. 2006. Fisheries: deep-sea fishes qualify as endangered. *Nature,* 439**,** 29.
Devine, J. A. & Haedrich, R. L. 2008. Population trends and status of two exploited Northwest Atlantic grenadiers, *Coryphaenoides rupestris* and *Macrourus berglax*. *American Fisheries Society Symposium,* 63**,** 463-484.
Drazen, J. 2002. A seasonal analysis of the nutritional condition of deep-sea macrourid fishes in the north-east Pacific. *Journal of Fish Biology,* 60**,** 1280-1295.
Drazen, J. C. & Seibel, B. A. 2007. Depth-related trends in metabolism of benthic and benthopelagic deep-sea fishes. *Limnology and Oceanography,* 52**,** 2306-2316.
Drazen, J. C. & Sutton, T. T. 2017. Dining in the Deep: The Feeding Ecology of Deep-Sea Fishes. *Annual Review of Marine Science,* 9**,** 337-366.
Ebeling, A. W. & Cailliet, G. M. 1974. Mouth size and predator strategy of midwater fishes. *Deep Sea Research and Oceanographic Abstracts,* 21**,** 959-968.
Emery, W. J. & Dewar, J. 1982. Mean temperature-salinity, salinity-depth and temperature-depth curves for the North Atlantic and the North Pacific. *Progress in Oceanography,* 11**,** 219-305.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. 2010. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography,* 66**,** 486-501.
Escuret, V., Collins, P. J., Casalegno, J. S., Vachieri, S. G., Cattle, N., Ferraris, O., Sabatier, M., Frobert, E., Caro, V., Skehel, J. J., Gamblin, S., Valla, F., Valette, M., Ottmann, M., Mccauley, J. W., Daniels, R. S. & Lina, B. 2014. A novel I221L substitution in neuraminidase confers high-level resistance to oseltamivir in influenza B viruses. *Journal of Infectious Diseases,* 210**,** 1260-9.
Foltz, G. R. & Mcphaden, M. J. 2008. Seasonal mixed layer salinity balance of the tropical North Atlantic Ocean. *Journal of Geophysical Research: Oceans,* 113.
French, A., Macedo, M., Poulsen, J., Waterson, T. & Yu, A. 2008. Multivariate analysis of variance (MANOVA). San Francisco State University.
Froese, R. 2006. Cube law, condition factor and weight–length relationships: history, meta-analysis and recommendations. *Journal of applied ichthyology,* 22**,** 241-253.
Fromm, J. A., Johnson, S. A. & Johnson, D. L. 2008. Epidermal growth factor receptor 1 (EGFR1) and its variant EGFRvIII regulate TATA-binding protein expression through distinct pathways. *Molecular and cellular biology,* 28**,** 6483-6495.
Fukuzawa, A., Lange, S., Holt, M., Vihola, A., Carmignac, V., Ferreiro, A., Udd, B. & Gautel, M. 2008. Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band: implications for hereditary myopathies. *Journal of Cell Science,* 121**,** 1841-51.
Gage, J. D. & Tyler, P. A. 1991. *Deep-sea biology: a natural history of organisms at the deep-sea floor*, Cambridge University Press.
Gaither, M. R., Bowen, B. W., Rocha, L. A. & Briggs, J. C. 2016. Fishes that rule the world: circumtropical distributions revisited. *Fish and Fisheries,* 17**,** 664-679.
Gaither, M. R., Gkafas, G. A., De Jong, M., Sarigol, F., Neat, F., Regnier, T., Moore, D., Grcke, D. R., Hall, N., Liu, X., Kenny, J., Lucaci, A., Hughes, M., Haldenby, S. & Hoelzel, A. R. 2018. Genomics of habitat choice and adaptive evolution in a deep-sea fish. *Nature Ecology & Evolution,* 2**,** 680-687.
Gaither, M. R., Violi, B., Gray, H. W., Neat, F., Drazen, J. C., Grubbs, R. D., Roa-Varón, A., Sutton, T. & Hoelzel, A. R. 2016. Depth as a driver of evolution in the deep sea: Insights from grenadiers (Gadiformes: Macrouridae) of the genus Coryphaenoides. *Molecular Phylogenetics & Evolution,* 104**,** 73-82.
Gao, M., Lu, H. & Schulten, K. 2001. Simulated refolding of stretched titin immunoglobulin domains. *Biophysical Journal,* 81**,** 2268-77.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research,* 31**,** 3784-8.
Geisler, S. B., Robinson, D., Hauringa, M., Raeker, M. O., Borisov, A. B., Westfall, M. V. & Russell, M. W. 2007. Obscurin-like 1, OBSL1, is a novel cytoskeletal protein related to obscurin. *Genomics,* 89**,** 521-31.
Gerday, C., Aittaleb, M., Arpigny, J. L., Baise, E., Chessa, J. P., Garsoux, G., Petrescu, I. & Feller, G. 1997. Psychrophilic enzymes: a thermodynamic challenge. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology,* 1342**,** 119-31.
Gillett, M. B., Suko, J. R., Santoso, F. O. & Yancey, P. H. 1997. Elevated levels of trimethylamine oxide in muscles of deep-sea gadiform teleosts: A high-pressure adaptation? *Journal of Experimental Zoology,* 279**,** 386-391.
Glover, A., Smith, C., Paterson, G., Wilson, G., Hawkins, L. & Sheader, M. 2002. Polychaete species diversity in the central Pacific abyss: local and regional patterns, and relationships with productivity. *Marine Ecology Progress Series,* 240**,** 157-170.
Gordon, J. & Swan, S. 1996. Validation of age readings from otoliths of juvenile roundnose grenadier, *Coryphaenoides rupestris*, a deep-water macrourid fish. *Journal of Fish Biology,* 49**,** 289-297.
Gordon, J., Swan, S., Kelly, C. & Hareide, N. 1995. Age determination of juvenile roundnose grenadier *Coryphaenoides rupestris*, a deep water macrourid fish: a preliminary report. *ICES CM.*
Grabowski, E., Letelier, R. M., Laws, E. A. & Karl, D. M. 2019. Coupling carbon and energy fluxes in the North Pacific Subtropical Gyre. *Nature Communications,* 10**,** 1895.
Gross, M. & Jaenicke, R. 1994. Proteins under pressure: The influence of high hydrostatic pressure on structure, function and assembly of proteins and protein complexes. *European Journal of Biochemistry,* 221**,** 617-30.
Haedrich, R., Merrett, N. & O'dea, N. 2001. Can ecological knowledge catch up with deep-water fishing? A North Atlantic perspective. *Fisheries Research,* 51**,** 113-122.
Hanson, D., Murray, P. G., Coulson, T., Sud, A., Omokanye, A., Stratta, E., Sakhinia, F., Bonshek, C., Wilson, L. C., Wakeling, E., Temtamy, S. A., Aglan, M., Rosser, E. M., Mansour, S., Carcavilla, A., Nampoothiri, S., Khan, W. I., Banerjee, I., Chandler, K. E., Black, G. C. & Clayton, P. E. 2012. Mutations in CUL7, OBSL1 and CCDC8 in 3-M syndrome lead to disordered growth factor signalling. *Journal of Molecular Endocrinology,* 49**,** 267-75.
Hardy, F., Vriend, G., Veltman, O. R., Van Der Vinne, B., Venema, G. & Eijsink, V. G. 1993. Stabilization of *Bacillus stearothermophilus* neutral protease by introduction of prolines. *FEBS Letters,* 317**,** 89-92.
Hessler, R. R. & Sanders, H. L. 1967. Faunal diversity in the deep-sea. *Deep Sea Research and Oceanographic Abstracts,* 14**,** 65-78.
Huber, C., Fradin, M., Edouard, T., Le Merrer, M., Alanay, Y., Da Silva, D. B., David, A., Hamamy, H., Van Hest, L., Lund, A. M., Michaud, J., Oley, C., Patel, C., Rajab, A., Skidmore, D. L., Stewart, H., Tauber, M., Munnich, A. & Cormier-Daire, V. 2010. OBSL1 mutations in 3-M syndrome are associated with a modulation of IGFBP2 and IGFBP5 expression levels. *Human Mutations,* 31**,** 20-6.
Hull, J., Campino, S., Rowlands, K., Chan, M.-S., Copley, R. R., Taylor, M. S., Rockett, K., Elvidge, G., Keating, B. & Knight, J. 2007. Identification of common genetic variation that modulates alternative splicing. *PLoS Genetics,* 3**,** e99.
Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M. & Kimchi-Sarfaty, C. 2009. Silent (synonymous) SNPs: should we care about them? *In:* Komar, A. A. (ed.) *Single Nucleotide Polymorphisms.* Springer.
Ices 2016. ICES Advice on fishing opportunities, catch, and effort, Greater North Sea Ecoregion. *ICES Advice 2016, Book 9.*
.
Iwamoto, T. 2015. *Coryphaenoides rupestris* [Online]. Available: http://dx.doi.org/10.2305/IUCN.UK.2015-4.RLTS.T15522149A15603540.en. [Accessed 19 July 2019].
Jagodzinski, F., Hardy, J. & Streinu, I. 2012. Using rigidity analysis to probe mutation-induced structural changes in proteins. *Journal of Bioinformatics & Computational Biology,* 10**,** 1242010.
Jewell, P. W. 1994. Deep water nutrient and oxygen gradients in a modern coastal upwelling zone and their paleoceanographic implications. *Journal of Geophysical Research: Oceans,* 99**,** 7845-7850.
Kaariainen, J. & Bett, B. 2006. Evidence for benthic body size miniaturization in the deep sea. *Journal of the Marine Biological Association of the UK,* 86**,** 1339-1345.
Kim, H. Y. 2013. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics,* 38**,** 52-4.
Kon, T., Oyama, T., Shimo-Kon, R., Imamula, K., Shima, T., Sutoh, K. & Kurisu, G. 2012. The 2.8 A crystal structure of the dynein motor domain. *Nature,* 484**,** 345-50.

Koslow, J. A., Boehlert, G., Gordon, J., Haedrich, R., Lorance, P. & Parin, N. 2000. Continental slope and deep-sea fisheries: implications for a fragile ecosystem. *ICES Journal of Marine Science,* 57**,** 548-557.
Lambshead, P. J. D., Brown, C. J., Ferrero, T. J., Mitchell, N. J., Smith, C. R., Hawkins, L. E. & Tietjen, J. 2002. Latitudinal diversity patterns of deep-sea marine nematodes and organic fluxes: a test from the central equatorial Pacific. *Marine Ecology Progress Series,* 236**,** 129-135.
Lange, S., Pinotsis, N., Agarkova, I. & Ehler, E. 2020. The M-band: The underestimated part of the sarcomere. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research,* 1867**,** 118440.
Levin, L. A., Etter, R. J., Rex, M. A., Gooday, A. J., Smith, C. R., Pineda, J., Stuart, C. T., Hessler, R. R. & Pawson, D. 2001. Environmental influences on regional deep-sea species diversity. *Annual Review of Ecology & Systematics,* 32**,** 51-93.
Likhachev, E. 2003. Dependence of Water Viscosity on Temperature and Pressure. *Technical Physics,* 48.
Lim, Y., Yoo, J., Kim, M. S., Hur, M., Lee, E. H., Hur, H. S., Lee, J. C., Lee, S. N., Park, T. W., Lee, K., Chang, K. H., Kim, K., Kang, Y., Hong, K. W., Kim, S. H., Kim, Y. G., Yoon, Y., Nam, D. H., Yang, H., Kim, D. G., Cho, H. S. & Won, J. 2016. GC1118, an Anti-EGFR Antibody with a Distinct Binding Epitope and Superior Inhibitory Activity against High-Affinity EGFR Ligands. *Molecular Cancer Therapeutics,* 15**,** 251-63.
Lin, H.-Y., Shiao, J.-C., Chen, Y.-G. & Iizuka, Y. 2012. Ontogenetic vertical migration of grenadiers revealed by otolith microstructures and stable isotopic composition. *Deep Sea Research Part I: Oceanographic Research Papers,* 61**,** 123-130.
Linley, T. D., Gerringer, M. E., Yancey, P. H., Drazen, J. C., Weinstock, C. L. & Jamieson, A. J. 2016. Fishes of the hadal zone including new species, in situ observations and depth records of *Liparidae*. *Deep Sea Research Part I: Oceanographic Research Papers,* 114**,** 99-110.
Longmore, C., Trueman, C., Neat, F., O'gorman, E., Milton, J. & Mariani, S. 2011. Otolith geochemistry indicates life-long spatial population structuring in a deep-sea fish, *Coryphaenoides rupestris*. *Marine Ecology Progress Series,* 435**,** 209-224.
Lorance, P., Dupouy, H. & Allain, V. 2001. Assessment of the roundnose grenadier (*Coryphaenoides rupestris*) stock in the Rockall Trough and neighbouring areas (ICES Sub-areas V–VII). *Fisheries Research,* 51**,** 151-163.
Lorance, P., Garren, F. & Vigneau, J. 2003. Age estimation of roundnose grenadier (*Coryphaenoides rupestris*), effects of uncertainties on ages. *Journal of Northwest Atlantic Fishery Science,* 31.
Lorance, P. & Trenkel, V. M. 2006. Variability in natural behaviour, and observed reactions to an ROV, by mid-slope fish species. *Journal of Experimental Marine Biology and Ecology,* 332**,** 106-119.
Lu, H., Isralewitz, B., Krammer, A., Vogel, V. & Schulten, K. 1998. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophysical Journal,* 75**,** 662-71.
Madden, T. 2013. The BLAST sequence analysis tool. *The NCBI Handbook.* 2nd ed.: National Center for Biotechnology Information (US).
Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D. & Lopez, R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research,* 47**,** W636-W641.
Magnuson, J. & Heitz, J. G. 1971. Gill raker apparatus and food selectivity among mackerels, tunas, and dolphins. *Fish Bull,* 69**,** 361-370.
Mauchline, J., Bergstad, O. A., Gordon, J. D. & Brattegard, T. 1994. The food of juvenile *Coryphaenoides rupestris* Gunnerus, 1765 (Pisces, Macrouridae) in the Skagerrak. *Sarsia,* 79**,** 163-164.
Mauchline, J. & Gordon, J. D. 1991. Oceanic pelagic prey of benthopelagic fish in the benthic boundary layer of a marginal oceanic region. *Marine Ecology Progress Series,* 74**,** 109-115.
Mcdonald, J. H. 2009. *Handbook of Biological Statistics,* Baltimore, MD, Sparky House Publishing
Morales-Nin, B. & Panfili, J. 2005. Seasonality in the deep sea and tropics revisited: what can otoliths tell us? *Marine and Freshwater Research,* 56**,** 585-598.
Morita, T. 2003. Structure-based analysis of high pressure adaptation of alpha-actin. *Journal of Biological Chemistry,* 278**,** 28060-6.
Musick, J. & Cotton, C. 2015. Bathymetric limits of chondrichthyans in the deep sea: a re-evaluation. *Deep Sea Research Part II: Topical Studies in Oceanography,* 115**,** 73-80.
Neat, F. C. 2017. Aggregating behaviour, social interactions and possible spawning in the deep-water fish Coryphaenoides rupestris. *Journal of Fish Biology,* 91**,** 975-980.
Neat, F. C. & Burns, F. 2010. Stable abundance, but changing size structure in grenadier fishes (Macrouridae) over a decade (1998–2008) in which deepwater fisheries became regulated. *Deep Sea Research Part I: Oceanographic Research Papers,* 57**,** 434-440.
Neat, F. C. & Campbell, N. 2013. Proliferation of elongate fishes in the deep sea. *Journal of Fish Biology,* 83**,** 1576-91.

Niitsu, A., Egawa, A., Ikeda, K., Tachibana, K. & Fujiwara, T. 2018. Veratridine binding to a transmembrane helix of sodium channel Nav1.4 determined by solid-state NMR. *Bioorganic & Medicinal Chemistry,* 26**,** 5644-5653.
Niven, J. E. 2015. Neural Evolution: Costing the Benefits of Eye Loss. *Current Biology,* 25**,** R840-1.
O'hea, B., Johnston, G., White, J. & Dransfeld, L. 2013. Length-weight relations for seven grenadier species (Actinopterygii: Gadiformes: Macrouridae) to the west of Ireland. *ACTA Ichthyologica et Piscatoria,* 43**,** 285-291.
Pelster, B. 1997. Buoyancy at Depth. *In:* Randall, D. J. F., Anthony P. (ed.) *Deep-Sea Fishes.* Academic Press.
Pelster, B. & Weber, R. E. 1991. The physiology of the Root effect. *Advances in Comparative and Environmental Physiology.* Springer.
Pernigo, S., Fukuzawa, A., Bertz, M., Holt, M., Rief, M., Steiner, R. A. & Gautel, M. 2010. Structural insight into M-band assembly and mechanics from the titin-obscurin-like-1 complex. *Proceedings of the National Academy of Sciences,* 107**,** 2908-2913.
Peters, R. H. 1986. *The ecological implications of body size*, Cambridge University Press.
Phleger, C. F. 1991. Biochemical aspects of buoyancy in fishes. *In:* Hochachka, P. W. a. M., T.P. (ed.) *Biochemistry and molecular biology of fishes.* Elsevier.
Phleger, C. F. 1998. Buoyancy in marine fishes: direct and indirect role of lipids. *American Zoologist,* 38**,** 321-330.
Porter, M. L., Roberts, N. W. & Partridge, J. C. 2016. Evolution under pressure and the adaptation of visual pigment compressibility in deep-sea environments. *Molecular Phylogenetics & Evolution,* 105**,** 160-165.
Priede, I. G. 2017. *Deep-sea fishes: biology, diversity, ecology and fisheries,* Cambridge University Press.
Priede, I. G., Froese, R., Bailey, D. M., Bergstad, O. A., Collins, M. A., Dyb, J. E., Henriques, C., Jones, E. G. & King, N. 2006. The absence of sharks from abyssal regions of the world's oceans. *Proceedings of the Royal Society B: Biological Sciences,* 273**,** 1435-41.
Priede, I. G., Ragley, P. & Smith Jr, K. 1994. Seasonal change in activity of abyssal demersal scavenging grenadiers *Coryphaenoides (Nematonums) armatus* in the eastern North Pacific Ocean. *Limnology and Oceanography,* 39**,** 279-285.
Priede, I. G., Smith Jr, K. L. & Armstrong, J. D. 1990. Foraging behavior of abyssal grenadier fish: inferences from acoustic tagging and tracking in the North Pacific Ocean. *Deep Sea Research Part A Oceanographic Research Papers,* 37**,** 81-101.
Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C. R., Levin, L. A., Arbizu, P. M., Menot, L. & Buhl-Mortensen, P. 2010. Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. *Biogeosciences,* 7**,** 2851-2899.
Rex, M. A. & Etter, R. J. 1998. Bathymetric patterns of body size: implications for deep-sea biodiversity. *Deep-Sea Research Part II: Topical Studies in Oceanography,* 45**,** 103-127.
Rex, M. A., Etter, R. J., Morris, J. S., Crouse, J., Mcclain, C. R., Johnson, N. A., Stuart, C. T., Deming, J. W., Thies, R. & Avery, R. 2006. Global bathymetric patterns of standing stock and body size in the deep-sea benthos. *Marine Ecology Progress Series,* 317**,** 1-8.
Rex, M. A., Stuart, C. T. & Coyne, G. 2000. Latitudinal gradients of species richness in the deep-sea benthos of the North Atlantic. *Proceedings of the National Academy of Sciences,* 97**,** 4082-4085.
Rex, M. A., Stuart, C. T., Hessler, R. R., Allen, J. A., Sanders, H. L. & Wilson, G. D. 1993. Global-scale latitudinal patterns of species diversity in the deep-sea benthos. *Nature,* 365**,** 636-639.
Reygondeau, G., Guidi, L., Beaugrand, G., Henson, S. A., Koubbi, P., Mackenzie, B. R., Sutton, T. T., Fioroni, M. & Maury, O. 2018. Global biogeochemical provinces of the mesopelagic zone. *Journal of Biogeography,* 45**,** 500-514.
Rios, F. S. A., Moraes, G., Oba, E. T., Fernandes, M. N., Donatti, L., Kalinin, A. L. & Rantin, F. T. 2006. Mobilization and recovery of energy stores in traíra, *Hoplias malabaricus* Bloch (Teleostei, Erythrinidae) during long-term starvation and after re-feeding. *Journal of Comparative Physiology B,* 176**,** 721-728.
Rogers, A. D. 2000. The role of the oceanic oxygen minima in generating biodiversity in the deep sea. *Deep Sea Research Part II: Topical Studies in Oceanography,* 47**,** 119-148.
Ryther, J. H. 1956. Photosynthesis in the Ocean as a Function of Light Intensity. *Limnology and Oceanography,* 1**,** 61-70.
Salvanes, A. & Kristoffersen, J. 2001. Mesopelagic Fishes. *In:* Steele, J. H. (ed.) *Marine Biology.* Academic Press.
Sauer, F. 2011. *Structural studies on the association of filamentous proteins in the human M-Bands.* PhD, University of Würzburg.
Savvatimsky, P. 1982. Reproduction and sex composition of the North Atlantic roundnose grenadier. *Abundance and modes of life of the Northwest Atlantic commercial fishes Canadian Transactions of Fisheries and Aquatic Sciences 5389.*

Savvatimsky, P. 1985. On correlation between total length and preanal length of roundnose grenadier in the North Atlantic. *NAFO Scientific Council Studies,* 8**,** 61-65.
Schneider, J. C., Laarman, P. W., Howard Gowing, J. C., Laarman, P. & Gowing, H. 2000. Length-weight relationships. *In:* Schneider, J. C. (ed.) *Manual of fisheries survey methods II: With periodic updates.* Michigan Department of Natural Resources.
Schoenauer, R., Bertoncini, P., Machaidze, G., Aebi, U., Perriard, J.-C., Hegner, M. & Agarkova, I. 2005. Myomesin is a molecular spring with adaptable elasticity. *Journal of Molecular Biology,* 349**,** 367-379.
Scholander, P. F. & Van Dam, L. 1954. Secretion of gases against high pressures in the swimbladder of deep sea fishes I. oxygen dissociation in blood. *The Biological Bulletin,* 107**,** 247-259.
Shum, P., Pampoulie, C., Sacchi, C. & Mariani, S. 2014. Divergence by depth in an oceanic fish. *PeerJ,* 2**,** e525.
Siebenaller, J. F. & Somero, G. N. 1982. The maintenance of different enzyme activity levels in congeneric fishes living at different depths. *Physiological Zoology,* 55**,** 171-179.
Siebenaller, J. F., Somero, G. N. & Haedrich, R. L. 1982. Biochemical characteristics of macrourid fishes differing in their depths of distribution. *The Biological Bulletin,* 163**,** 240-249.
Simpson, M., Miri, C., Mercer, J., Bailey, J., Power, D. J., Themelis, D. & Treble, M. 2011. Recovery Potential Assessment for Roundnose Grenadier (Coryphaenoides Rupestris Gunnerus, 1765) in Northwest Atlantic Waters. Fisheries and Oceans Canada, Science, Newfoundland & Labrador Region.
Sivasundar, A. & Palumbi, S. R. 2010. Parallel amino acid replacements in the rhodopsins of the rockfishes (*Sebastes spp.*) associated with shifts in habitat depth. *Journal of Evolutionary Biology,* 23**,** 1159-69.
Slinker, B. & Glantz, S. 1985. Multiple regression for physiological data analysis: the problem of multicollinearity. *American Journal of Physiology-Regulatory, Integrative & Comparative Physiology,* 249**,** R1-R12.
Smith, C. R., Levin, L. A., Koslow, A., Tyler, P. A. & Glover, A. G. 2008. The near future of the deep seafloor ecosystems. *In:* Polunin, N. (ed.) *Aquatic Ecosystems: Trends and Global Prospects.* Cambrige University Press.
Smith, K. 1978. Metabolism of the abyssopelagic rattail *Coryphaenoides armatus* measured in situ. *Nature,* 274**,** 362-364.
Solovyev, V., Kosarev, P., Seledsov, I. & Vorobyev, D. 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology,* 7 Suppl 1**,** S10 1-12.
Somero, G. 1998. Adaptation to cold and depth: contrasts between polar and deep-sea animals. *In:* Portner, H.-O. P., R.C. (ed.) *Cold ocean physiology.* Cambridge University Press.
Somero, G. N. 1990. Life at low volume change: hydrostatic pressure as a selective factor in the aquatic environment. *American Zoologist,* 30**,** 123-135.
Somero, G. N. & Siebenaller, J. F. 1979. Inefficient lactate dehydrogenases of deep-sea fishes. *Nature,* 282**,** 100-2.
Somkuti, J., Martonfalvi, Z., Kellermayer, M. S. & Smeller, L. 2013. Different pressure-temperature behavior of the structured and unstructured regions of titin. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics,* 1834**,** 112-118.
Squire, J. M., Al-Khayat, H. A., Knupp, C. & Luther, P. K. 2005. Molecular architecture in muscle contractile assemblies. *Advances in Protein Chemistry,* 71**,** 17-87.
Sullivan, K. M. & Somero, G. N. 1983. Size-and diet-related variations in enzymic activity and tissue composition in the sablefish, *Anoplopoma fimbria*. *The Biological Bulletin,* 164**,** 315-326.
Sutton, T. T., Wiebe, P. H., Madin, L. & Bucklin, A. 2010. Diversity and community structure of pelagic fishes to 5000 m depth in the Sargasso Sea. *Deep Sea Research Part II: Topical Studies in Oceanography,* 57**,** 2220-2233.
Swezey, R. R. & Somero, G. N. 1985. Pressure effects on actin self-assembly: interspecific differences in the equilibrium and kinetics of the G to F transformation. *Biochemistry,* 24**,** 852-60.
Thiel, H. 1975. The size structure of the deep-sea benthos. *Internationale Revue der Gesamten Hydrobiologie,* 60**,** 575-606.
Thiel, H. 1979. Structural aspects of the deep-sea benthos. *Ambio Special Report***,** 25-31.
Thurber, A. R., Sweetman, A. K., Narayanaswamy, B. E., Jones, D. O., Ingels, J. & Hansman, R. 2014. Ecosystem function and services provided by the deep sea. *Biogeosciences,* 11**,** 3941-3963.
Ungermann, C. & Reggiori, F. 2018. Atg9 proteins, not so different after all. *Autophagy,* 14**,** 1456-1459.
Van Dover, C. L., German, C., Speer, K. G., Parson, L. & Vrijenhoek, R. 2002. Evolution and biogeography of deep-sea vent and seep invertebrates. *Science,* 295**,** 1253-1257.
Van Rees, W. M., Gazzola, M. & Koumoutsakos, P. 2013. Optimal shapes for anguilliform swimmers at intermediate Reynolds numbers. *Journal of Fluid Mechanics,* 722.

Vetter, R., Lynn, E., Garza, M. & Costa, A. 1994. Depth zonation and metabolic adaptation in Dover sole, *Microstomus pacificus*, and other deep-living flatfishes: factors that affect the sole. *Marine Biology,* 120**,** 145-159.
Vetter, R. D. & Lynn, E. A. 1997. Bathymetric demography, enzyme activity patterns, and bioenergetics of deep-living scorpaenid fishes (genera *Sebastes* and *Sebastolobus*): paradigms revisited. *Marine Ecology Progress Series,* 155**,** 173-188.
Wakai, N., Takemura, K., Morita, T. & Kitao, A. 2014. Mechanism of deep-sea fish alpha-actin pressure tolerance investigated by molecular dynamics simulations. *PLoS One,* 9**,** e85852.
Warrant, E. J. 2000. The eyes of deep–sea fishes and the changing nature of visual scenes with depth. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences,* 355**,** 1155-1159.
Warrant, E. J., Collin, S. P. & Locket, N. A. 2003. Eye design and vision in deep-sea fishes. *In:* Collin, S. P. M., N.J. (ed.) *Sensory Processing in Aquatic Environments.* Springer.
White, T. A., Stamford, J. & Rus Hoelzel, A. 2010. Local selection and population structure in a deep-sea fish, the roundnose grenadier (*Coryphaenoides rupestris*). *Molecular Ecology,* 19**,** 216-226.
Wilson, G. D. & Ahyong, S. T. 2015. Lifestyles of the species-rich and fabulous: the deep-sea crustaceans. *In:* Thiel, M. W., L (ed.) *Life Styles and Feeding Biology.* Oxford University Press.
Wilson, G. D. & Hessler, R. R. 1987. Speciation in the deep sea. *Annual review of ecology and systematics,* 18**,** 185-207.
Wood, E. R., Shewchuk, L. M., Ellis, B., Brignola, P., Brashear, R. L., Caferro, T. R., Dickerson, S. H., Dickson, H. D., Donaldson, K. H. & Gaul, M. 2008. 6-Ethynylthieno [3, 2-d]-and 6-ethynylthieno [2, 3-d] pyrimidin-4-anilines as tunable covalent modifiers of ErbB kinases. *Proceedings of the National Academy of Sciences,* 105**,** 2773-2778.
Wu, H., Wang, C., Gong, W., Wang, J., Xuan, J., Perrett, S. & Feng, Y. 2016. The C-terminal region of human eukaryotic elongation factor 1Bdelta. *Journal of Biomolecular NMR,* 64**,** 181-7.
Wyrtki, K. 1962. The oxygen minima in relation to ocean circulation. *Deep Sea Research and Oceanographic Abstracts,* 9**,** 11-23.
Yamaguchi, H., Kasa, M., Amano, M., Kaibuchi, K. & Hakoshima, T. 2006. Molecular mechanism for the regulation of rho-kinase by dimerization and its inhibition by fasudil. *Structure,* 14**,** 589-600.
Yancey, P. H., Blake, W. R. & Conley, J. 2002. Unusual organic osmolytes in deep-sea animals: adaptations to hydrostatic pressure and other perturbants. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology,* 133**,** 667-76.
Yancey, P. H., Fyfe-Johnson, A. L., Kelly, R. H., Walker, V. P. & Aunon, M. T. 2001. Trimethylamine oxide counteracts effects of hydrostatic pressure on proteins of deep-sea teleosts. *Journal of Experimental Zoology,* 289**,** 172-6.
Yancey, P. H., Rhea, M. D., Kemp, K. M. & Bailey, D. M. 2004. Trimethylamine oxide, betaine and other osmolytes in deep-sea animals: depth trends and effects on enzymes under hydrostatic pressure. *Cellular and Molecular Biology,* 50**,** 371-6.
Yang, T.-H. & Somero, G. N. 1993. Effects of feeding and food deprivation on oxygen consumption, muscle protein concentration and activities of energy metabolism enzymes in muscle and brain of shallow-living (*Scorpaena guttata*) and deep-living (*Sebastolobus alascanus*) scorpaenid fishes. *Journal of Experimental Biology,* 181**,** 213-232.
Zou, P., Pinotsis, N., Lange, S., Song, Y. H., Popov, A., Mavridis, I., Mayans, O. M., Gautel, M. & Wilmanns, M. 2006. Palindromic assembly of the giant muscle protein titin in the sarcomeric Z-disk. *Nature,* 439**,** 229-33.

# Webography

EMBOSS Water Pairwise Alignment: https://www.ebi.ac.uk/Tools/psa/emboss_water/

ExPASY Translate: https://web.expasy.org/translate/

Fgenesh: http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind

GeneWise Protein-Gene Pair Alignment: https://www.ebi.ac.uk/Tools/psa/genewise/

JBrowse: https://jbrowse.org/

ModWeb: https://modbase.compbio.ucsf.edu/modweb/

NCBI BLAST: https://blast.ncbi.nlm.nih.gov/Blast.cgi

PDBe: https://www.ebi.ac.uk/pdbe/node/1

UniProt: https://www.uniprot.org/

UniProt BLAST: https://www.uniprot.org/blast/