

Genetic and morphometric variation of two focal
snapper species (*Lutjanus bohar* and *Lutjanus gibbus*)
in the Chagos Archipelago

Abigail S. Bailey

A thesis submitted in partial fulfilment of the requirements of Oxford Brookes University for the
degree of Doctor of Philosophy

Faculty of Health and Life Sciences
Oxford Brookes University

November 2021

Abstract

Genetic and morphological variation were investigated for two snapper species: two-spotted snapper (*Lutjanus bohar*) and humpback red snapper (*Lutjanus gibbus*), at four atolls (Diego Garcia, Egmont, Peros Banhos and Salomon) within the Chagos Archipelago in the Indian Ocean. This is an area which has some of the most undisturbed and remote corals on earth, and an ideal location for understanding how species within reef communities might adapt to environmental change and degradation.

Three separate approaches were used to carry out this investigation:

43 individual *L. bohar* were screened for genetic variation at the cytochrome c oxidase subunit 1 (CO1) mitochondrial DNA (mtDNA) gene. Results show overall high mtDNA diversity for the populations, but low nucleotide divergence between haplotypes. The species constitutes a well-mixed population with a ubiquitous, predominant haplotype, shared by 51% of specimens.

129 *L. bohar* and 101 *L. gibbus* individuals were assessed for morphometric (size and shape) variability. A significant difference in mean size for both lutjanid species was detected between Diego Garcia and each of the three other atolls. For both species, smallest individuals were found at Diego Garcia, where recreational fishing has depleted fish biomass in general and may have also directly impacted the abundance of large lutjanids. Variations in shape (weight, circumference and body depth, relative to fork length) between locations were not statistically significant when collinearity was accounted for.

118 *L. bohar* and 69 *L. gibbus* were selected for population structure analysis using a next-generation sequencing RAD-seq approach. A genomic analysis of population structure was carried out using genome wide SNP markers. Weak population structure was detected between all atolls, for both species (F_{st} 0.001-0.009) indicating good genetic connectivity among lutjanids throughout the Chagos Archipelago.

The combined findings of this research are consistent with the hypothesis that there is little population structure for *L. bohar* and *L. gibbus* within the Chagos Archipelago and that populations are well connected. As a result, both species are likely to be genetically homogenous throughout the geographic region. The fact that morphometric variation between atolls was restricted to differences in size indicates a plastic response to food availability rather than an evolutionary adaptive response.

This research represents the first comparison of genetic and morphometric variation amongst two focal snapper species in the Chagos Archipelago.

Acknowledgements

I owe a wealth of thanks to my husband Neil, and my two children Safi and Jack who have become adults while I've been the eternal student. This work would not have been possible without my family's constant support, Safi's amazing proof-reading, and Jack's grounding critique. Without you all telling me I can finish this; I wouldn't have gotten to this point.

Thank you to my Director of Studies, Professor Stewart Thompson, for your constant support and encouragement throughout the years – the delivery of this project would not have been possible without your understanding and trust in me. A huge thank you is also due to my supervisor, Dr Michelle Taylor who introduced me to this project and has been there with me every step of the way – thank you for inspiring a love of coral reefs and everything marine, and for everything you have done to support me.

I am incredibly grateful to Jessica Meeuwig at UWA, for trusting me with her valuable samples, allowing me the opportunity to work on this exciting project, and for providing me with the lutjanid and shark data. I would also like to acknowledge the financial support of the UWA, who provided the funding for the genomic analysis of the samples, which enabled me to undertake this research.

I would also like to say thank you to the friends and colleagues who have made major contributions to this project including Dave Hemprich-Bennett – for the ongoing support and positivity as well as for helping me with R and helping me realise that it is not always terrifying and can be very useful – and to Jon Emberson for the statistics knowledge! I am truly grateful. And last, but very definitely not least, the support of my amazing friends has been fundamental to my PhD journey – Barbara, Marita, Andrea, Lucy and Kate – I'd be lost without you all. And big thanks to Nicky C for always being on my case!

Finally, I would like to dedicate this thesis to two people: my mum, Julia, who has shown us all how determination wins when faced with the impossible, since recovering from a devastating stroke in 2017; Mum, you make me proud every day. I would also like to remember my father-in-law, who was taken from us by COVID earlier this year. Graham: this is about fish and you will always be 'The Fisherman'.

We miss you.

Fishes live in the sea, as men do a-land; the great ones eat up the little ones.

William Shakespeare

*The sea, the great unifier, is man's only hope. Now, as never before,
the old phrase has a literal meaning: We are all in the same boat.*

Jacques Cousteau

*By the deep Sea, and music in its roar:
I love not Man the less, but Nature more.*

George Gordon Byron

Table of Contents

	Page no.
Chapter 1 General Introduction	1
1.1 Wildlife decline and threats to marine habitats	1
1.2 Morphometric variation - adaptive response and phenotypic diversity	2
1.3 Population genetics of fish in marine ecosystems	5
1.4 Detecting population structure in marine populations	7
1.4.1 Mitochondrial DNA (mtDNA) sequencing for phylogeographic analysis	8
1.4.2 Restriction Site Associated DNA Sequencing and related genotype by sequencing	8
1.5 The study area – the Chagos Archipelago Marine Protected Area	10
1.6 Investigating genetic and morphometric variation in two focal snapper species	12
1.6.1 <i>Lutjanus bohar</i>	13
1.6.2 <i>Lutjanus gibbus</i>	15
1.7 Aims	16
 Chapter 2 Variability in the CO1 Gene in Populations of <i>L. bohar</i> in the Chagos Archipelago	 17
2.1 Introduction	17
2.2 Methods	18
2.2.1 Sample Collection	18
2.2.2 DNA extraction	19
2.2.3 CO1 sequencing	19
2.2.4 Sequencing alignment	20
2.2.5 Haplotype network construction	20
2.2.6 Neutrality tests	20
2.3 Results	21
2.3.1 Sequence alignment	21
2.3.2 Haplotype network	24
2.3.3 Neutrality tests	26
2.4 Discussion	26
 Chapter 3 A morphometric analysis of <i>L. bohar</i> and <i>L. gibbus</i> within the Chagos Archipelago	 28
3.1 Introduction	28
3.2 Method	29
3.2.1 Sample collection	29
3.2.2 Size analysis	31
3.2.3 Shape analysis	32

3.3	Results – size variation	33
3.3.1	Summary size statistics for <i>L. bohar</i> and <i>L. gibbus</i>	33
3.3.2	ANOVA model to test for differences in size between sites for <i>L. bohar</i>	34
3.3.3	Post hoc Tukey tests to test for differences in size of <i>L. bohar</i> between sites	35
3.3.4	ANOVA model to test for differences in size between sites for <i>L. gibbus</i>	37
3.3.5	Post hoc Tukey tests to test for differences in size of <i>L. gibbus</i> between sites	38
3.3.6	Correlation test to measure the relationship between size variables	40
3.4	Results – shape variation	41
3.4.1	ANCOVA model to test or differences in shape of <i>L. bohar</i> between locations	41
3.4.2	Residual analysis for <i>L. bohar</i>	42
3.4.3	ANCOVA model to test for differences in shape of <i>L. gibbus</i> between locations	43
3.4.4	Residual analysis for <i>L. gibbus</i>	44
3.5	Discussion	45
3.5.1	Response to food availability	45
3.5.2	Direct fishing pressure	47
3.5.3	Predation risk	48
3.5.4	Habitat complexity	49
3.5.5	Temperature	50
3.6	Conclusion	52

Chapter 4 Population Genetic Structure of *L. bohar* and *L. gibbus* in the Chagos Archipelago **54**

4.1	Introduction	54
4.2	Methods	56
4.2.1	Sample collection	56
4.2.2	DNA extraction	57
4.2.3	RAD-seq library preparation	58
4.2.4	RAD-tag data analysis	59
4.2.5	Demultiplexing of data	59
4.2.6	Parameter selection	59
4.2.7	Population analysis using <i>Stacks populations</i> bioinformatics pipeline	60
4.2.8	PCA analysis of population structure using <i>adegenet</i> for R	61
4.3	Results	62
4.3.1	RAD-seq analysis: Running <i>Stacks process_radtags</i> to demultiplex data	62
4.3.2	RAD-seq analysis: Genotyping datasets for population analysis	62

4.3.3	RAD-seq analysis: F_{ST} analysis of population structure	63
4.3.4	Population structure of <i>L. bohar</i> using pairwise F_{ST} comparison	64
4.3.5	Population structure of <i>L. gibbus</i> using pairwise F_{ST} comparison	65
4.3.6	PCA analysis to detect clustering of individuals and populations	66
4.4	Discussion	69
4.4.1	Population genetic structure of <i>L. bohar</i> and <i>L. gibbus</i> in the Chagos Archipelago	69
4.4.2	Factors influencing population structure, gene flow and connectivity between atolls in the Chagos Archipelago	71
Chapter 5 Concluding Comments		76
Appendix: Chapter 3		85
Appendix: Chapter 4		89
Bibliography		100

List of Figures	Page No
Figure 1.1	Map showing the location of the Chagos Archipelago MPA 11
Figure 1.2	Adult <i>Lutjanus bohar</i> 14
Figure 1.3	Juvenile <i>Lutjanus bohar</i> 14
Figure 1.4	Adult <i>Lutjanus gibbus</i> 15
Figure 2.1	Sampling locations for <i>L. bohar</i> at atolls within the Chagos MPA 18
Figure 2.2	Multiple sequence alignments and phylogenetic tree belonging to the mitochondrial CO1 gene of <i>L. bohar</i> 21
Figure 2.3	Phylogenetic neighbour-joining tree belonging to the mitochondrial CO1 gene of <i>L. bohar</i> . 22
Figure 2.4	Median-joining haplotype network for the CO1 gene in 43 individuals of <i>L. bohar</i> 23
Figure 2.5	Histogram showing relative distribution of haplotypes for the CO1 gene in 43 individuals of <i>L. bohar</i> 24
Figure 3.1	Maps showing the sampling locations in the Chagos MPA for <i>L. bohar</i> and <i>L. gibbus</i> 29
Figure 3.2	Box plot showing average log weight differs between sites for <i>L. bohar</i> 34
Figure 3.3	Box plot showing how average log circumference differs between sites for <i>L. bohar</i> 35
Figure 3.4	Box plot showing how average log body depth differs between sites for <i>L. bohar</i> 35
Figure 3.5	Box plot showing how average log fork length differs between sites for <i>L. bohar</i> 36
Figure 3.6	Box plot showing average log weight differs between sites for <i>L. gibbus</i> 37
Figure 3.7	Box plot showing how average log circumference differs between sites for <i>L. gibbus</i> 38
Figure 3.8	Box plot showing how average log body depth differs between sites for <i>L. bohar</i> 38
Figure 3.9	Box plot showing how average log fork length differs between sites for <i>L. bohar</i> 39

Figure 4.1	Sampling locations for <i>L. bohar</i> at atolls within the Chagos MPA	55
Figure 4.2	Sampling locations for <i>L. gibbus</i> at atolls within the Chagos MPA	56
Figure 4.3	<i>L. bohar</i> – a representation of population structure between atolls for loci present in >75% of individuals at all sample sites	64
Figure 4.4	<i>L. gibbus</i> – a representation of population structure between atolls for loci present in >75% of individuals at all sample sites	65
Figure 4.5	Principal component analysis of allele frequencies for neutral loci present in 75% of the <i>L. bohar</i> population	66
Figure 4.6	Principal component analysis of allele frequencies for outlier loci present in 75% of the <i>L. bohar</i> population	67
Figure 4.7	Principal component analysis of allele frequencies for neutral loci present in 75% of the <i>L. gibbus</i> population	67
Figure 4.8	Principal component analysis of allele frequencies for outlier loci present in 75% of the <i>L. gibbus</i> population.	68
Figure 5.1	Summary infographic of research	77

List of Tables	Page No
Table 2.1 Numbers of <i>L. bohar</i> used for CO1 sequence alignment at each sampling location in the Chagos Archipelago	21
Table 2.2 Genetic variation within sites based on a 660 bp sequence of the mitochondrial CO1 gene for <i>L. bohar</i>	24
Table 3.1 Numbers of samples (n) collected at each sampling location and used for morphometric analysis	30
Table 3.2 Mean + Standard Error and minimum and maximum values for weight (kg), fork length (cm), circumference (cm) and body depth (cm) for <i>L. bohar</i> at each sampling location	32
Table 3.3 Mean + Standard Error and minimum and maximum values for weight (kg), fork length(cm), circumference(cm) and body depth(cm) for <i>L. gibbus</i> at each location.	33
Table 3.4 ANOVA comparing mean size metrics between sampling locations for <i>L. bohar</i>	33
Table 3.5 ANOVA comparing mean size metrics between sampling locations for <i>L. gibbus</i> .	36
Table 3.6 Pearson's correlation analyses between log transformed variables for <i>L. bohar</i> .	40
Table 3.7 Pearson's correlation analyses between log transformed variables for <i>L. gibbus</i>	40
Table 3.8 Results for ANCOVA models controlling for the fork length in <i>L. bohar</i> at all sampling sites to determine whether there are differences in shape between locations.	41
Table 3.9 Residual analysis to evaluate the independent contribution of response variables to determine whether differences in weight, circumference and body depth were statistically significant for <i>L. bohar</i> when body depth was accounted for and collinearity effects removed	42
Table 3.10 Results for ANCOVA models controlling for the fork length in <i>L. gibbus</i> at all sampling sites to determine whether there are differences in shape between locations	43
Table 3.11 Residual analysis to independent contribution of response variables to determine whether differences in weight, circumference and body depth were statistically significant for <i>L. gibbus</i> when body depth was accounted for and collinearity effects removed	43
Table 4.1 Percentage retained reads for run 1 and 2 following demultiplexing of raw RAD- sequences using process_radtags	61

Table 4.2	Number of loci and SNPs present in at least 75 to 95% of <i>L. bohar</i> and <i>L. gibbus</i> at all four sampling locations	62
Table 4.3	Numbers of <i>L. bohar</i> and <i>L. gibbus</i> individuals (n) at each sampling site used for population structure analysis and calculating F_{ST}	63
Table 4.4	F_{ST} values for pairwise comparisons between all four sampling sites for SNPs present in >75% of individuals for <i>L. bohar</i>	63
Table 4.5	F_{ST} values for pairwise comparisons between all four sampling sites for SNPs present in >75% of individuals for <i>L. gibbus</i>	65

Acronyms and Abbreviations

BIOT	British Indian Overseas Territory
HWE	Hardy Weinberg Equilibrium
IUU	Illegal Unregulated Unreported
MANOVA	Multivariate Analysis of Variance
MPA	Marine Protected Area
mtDNA	Mitochondrial DNA
NGS	Next Generation Sequencing
PCA	Principal Component Analysis
PLD	Pelagic Larval Duration
RAD	Restriction site Associated Digest
SNP	Single Nucleotide Polymorphism

Chapter 1 – General Introduction

1.1 Wildlife decline and threats to marine habitats

We are currently experiencing a ‘Sixth Extinction’ with species being lost at a rate previously only seen in mass extinction events (Matzke *et al.*, 2011). The most recent Living Planet Index report describes the situation in unambiguous terms: our planet is facing an unprecedented decline in species as a consequence of the pressure that humans exert (WWF/ZSL, 2018; WWF/ZSL, 2020). With over 60% of wildlife populations being lost over the past 40 years, never has there been a more crucial time to investigate the potential effects of biodiversity and species loss (WWF/ZSL, 2018).

The primary drivers of species decline are habitat degradation and loss (Böhm *et al.*, 2013; Collen *et al.*, 2014; Dirzo *et al.*, 2014), overexploitation (Dirzo *et al.*, 2014), and climate change (Steffen *et al.*, 2015; Gaffney and Steffen, 2017), all of which have ominous consequences for the Earth’s ecosystems and biodiversity. Rainforests are decreasing (Nobre *et al.*, 2016), the loss of warm water coral reefs and the marine ecosystems they support is unsustainable (Harvell *et al.*, 2008) and the Arctic is on a trajectory to be ice-free in summer months by 2050 (Jahn, 2018).

Marine habitats and marine species in particular face an uncertain future. Overfishing, selective fishing, climate change, and pollution are threatening marine environments. Almost 6 billion tonnes of fish and invertebrates have been taken from the oceans since 1950 (Pauly and Zeller, 2016).

Coral reef ecosystems represent some of the most diverse and species-rich environments on Earth, and are home to 25% of marine species (Reaka-Kudla, Wilson and Wilson, 1997), giving rise to them being considered “tropical rainforests of the sea”. In addition to supporting a wealth of biodiversity, tropical coral reefs are also valuable to humans, providing additional ecosystem benefits such as fisheries, tourism, and coastal storm protection (Ferrario *et al.*, 2014). Rising global sea temperatures and ocean acidification due to increasing CO₂ emissions, have caused unprecedented levels of coral bleaching in recent years (Sully *et al.*, 2019). Coral reef ecosystems are dynamic environments with an ability to tolerate periodic, natural cycles of disturbance and restructure in a bottom-up recovery process operating over several decades (Moran, Bradbury and Reichelt, 1988; Ninio *et al.*, 2000; Harvell *et al.*, 2002; Graham, Nash and Kool, 2011; Pandolfi *et al.*, 2011). However, resilience of coral reefs to recover fully is increasingly restricted due to accelerated human-induced disturbances (ocean acidification and warming, climate change induced cyclone damage, over-exploitation of fish populations and benthic degradation), operating on a global scale (Sully *et al.*, 2019).

Consequently, warm water coral reefs have declined by 50% in the last 30 years (Van Hooidonk *et al.*, 2014) and scientists predict that at current rates of degradation, 90% of the world's coral will be lost by 2050 (Heron *et al.*, 2016). In addition to the negative effects of climate change, a concerning anthropogenic disturbance afflicting coral reefs is fishing-related depletion of top-order predators in tropical coral reef ecosystems (Jackson *et al.*, 2001; Baum and Myers, 2004; Ferretti *et al.*, 2010; Ward-Paige *et al.*, 2010).

It is difficult to predict the effects of the loss of this important habitat and impacts on coral reef dwelling species and other communities that co-exist in complex and dynamic food webs. Caution suggests that multiple disturbances to coral reefs could have significant implications for ecosystem function, structure and interactions of coral reef communities and measures should be implemented to manage improved resilience (Bellwood *et al.*, 2004).

1.2 Morphometric variation - adaptive response and phenotypic diversity

As environments change, divergent natural selection occurs, often generating or maintaining phenotypic diversity. Phenotypic plasticity is the adaptive response and potential of an organism to modify its morphology to respond to a temporally or spatially heterogeneous environment. It is defined as the ability of a genotype to produce one or more different phenotypes in response to alterations in the environment. Thus, the phenotype of an organism (which includes morphological, behavioural, physiological, and biological traits) is driven by the complex interactions between genes (and the genetic capacity to accommodate environmental differences) and environment.

Phenotypic plasticity and the acclimatisation of an organism in response to changes in the natural environment is typically reversible, allowing flexibility for species to acclimatise to heterogeneous environments. The term 'reaction norm' has been used to describe phenotypic plasticity and the varied phenotypes that arise from a single genotype in response to environmental differences. Phenotypic plasticity and variation on phenotype caused by environmental change is a selective trait, enabling organisms to adapt to variability in the environment, such that the induced phenotype is advantageous. Phenotypic diversity can derive from either genetic differentiation or phenotypic plasticity, with both drivers inducing an adaptive response (Langerhans *et al.*, 2004).

In an ecological context, phenotypic plasticity is a useful tool enabling organisms to adapt to environmental change and habitat differences that arise in a heterogeneous environment (Pigliucci,

2005; Kelly, Panhuis and Stoeckl, 2012). The heterogeneity of an environment is characterised by physical factors such as gradients of physical structure (temperature, light, refuge etc.) and ecological interactions, such as predation, herbivory, and competition. Phenotypic responses to environmental factors are expressed across a diversity of organisms and can occur across generational time frames, or in a single individual's lifetime (Agrawal, LaForsch and Tollrian, 1999; Young, Stanton and Christian, 2003).

Examples of phenotypic plasticity include predator avoidance (Arnett and Kinnison, 2016), altered timing of metamorphosis in amphibians (Denver, Mirhadi and Phillips, 1998), osmoregulation in fish (Sollid and Nilsson, 2006), changes in reproductive behaviour (Bretman, Gage and Chapman, 2011), and the movement of amphibious fish between water and air (Wright and Turko, 2016).

Phenotypic plasticity is induced by a variety of mechanisms, in response to abiotic or biotic environmental characteristics and it is mediated via enzymatic or chemical influences. Phenotypic changes can also be caused by epigenetic modifications of the genome (e.g. DNA methylation to inhibit/activate DNA expression (Roberts and Gavery, 2012)), parental influence, learnt behaviour and memory (Jonsson and Jonsson, 2014).

Previous studies on phenotypic plasticity and adaptive responses in fish have focussed on morphometrics (shape and body size) and behaviour (boldness, increased vigilance, aggression) (Wisenden *et al.*, 2011; Andrew and Marco, 2012; King *et al.*, 2013). Fish species exhibit a diversity of phenotypes in response to different environmental drivers with morphological differences arising because of this phenotypic plasticity (Eklöv and Svanbäck, 2006; Chapman, 2015).

Various environmental factors may be important for influencing phenotypic plasticity of fish and resulting morphological variations, including predation risk, food availability, prey biology, habitat complexity, temperature, light, and hydrodynamics (Kröger and Fernald, 1994; Borcherting and Magnhagen, 2008; Collin and Fumagalli, 2011; Davis *et al.*, 2012; Chapman, 2015).

Predator-induced phenotypic plasticity is well described in natural environments and prey have been demonstrated to modify phenotype in response to various predator-induced cues. As well as being behavioural, these changes may be physical in nature, (e.g. changes in prey size or shape), and may confer a fitness benefit associated with the predator's reduced ability to consume an individual (Dewitt, Robinson and Wilson, 2000), or may increase the prey's ability to escape (Langerhans *et al.*,

2004). Examples of behavioural anti-predator plastic responses include the altered duration of activities such as feeding (Chalfoun and Martin, 2010), or mating (Johnson and Basolo, 2003). Adaptive responses to predators are usually associated with a trade-off in environments lacking a predator (Palkovacs, Wasserman and Kinnison, 2011; Lacasse and Aubin-Horth, 2015), which enforces a selective premium on an individual's potential to produce the most appropriate phenotypic response for any given predatory risk.

Food availability has been shown to impact fish growth and size. When food is scarce or plentiful, fish have been shown to develop morphological differences related to food availability, such that body size metrics (jaw depth, body depth, fork length) change (Currens *et al.*, 1989; Abaad *et al.*, 2016; Jacobson *et al.*, 2018). Prey size also influences fish and differs between locations for species where there are differences in prey size. As prey size increases so too does the gape, fork length and body depth of the predatory piscivore (and vice-versa) (Hegrenes, 2001; Paul *et al.*, 2017).

Differences in habitat complexity can also drive within-species morphological variation with fish acquiring an optimal shape for foraging, predator avoidance and locomotion in a specific environment (Olsson and Eklöv, 2005). Fish of the same species may exhibit different morphologies in different environments, with streamlined bodies providing advantage for acquiring food in open pelagic sea and facilitating fast-starts from predators, and deeper bodied individuals best adapted for movement in complex and structured habitats (Webb, 1984; Eklöv and Svanbäck, 2006; Mihalitsis and Bellwood, 2019).

One of the key factors that influences fish phenotype is temperature. Fish are poikilothermic and therefore sensitive to changes in temperature. The ambient environmental temperature determines fish body temperature and changes in temperature affect the biochemical reactions that determine the rate of physiological processes (Jonsson and Jonsson, 2014). Temperature affects metabolic processes by regulating feeding behaviour, food intake, and digestion, by controlling the activity of endocrine factors and digestive enzymes, within optimal thermal ranges (Folkeadal *et al.*, 2012; Jaxion-Harm and Ladich, 2014; Madison *et al.*, 2015; Park *et al.*, 2016). These processes control how much energy is consumed, acquired, and distributed for processes such as growth, locomotion, and reproduction. The response to any deviation from the optimal thermal range is dependent on the degree of phenotypic plasticity of the fish and its ability to acclimatise to any increase / decrease in temperature. Under extreme thermal stress (whether that is extreme cold, or heat) these metabolic

and physiological processes may be inhibited, with subsequent implications for the health and survival of the fish.

As a result of a combination of factors, many marine species have been shown to exhibit morphological changes in body shape, size and growth rate in response to diet and feeding ecology, with these variations in phenotype being associated with prey size, food availability and temperature (Sarà, Favaloro and Mazzola, 1999; Hegrenes, 2001; Favaloro and Mazzola, 2003; Domenici *et al.*, 2008). This morphological variation may be the result of relative contributions of plastic response to diet / altered environmental pressure and heritable genetic variation derived from individual parents (Eklöv and Svanbäck, 2006). Processes driving morphological change may be evidence of the early stages of divergence and speciation, as the evolution of body shape is a strong driving force that can push evolution in new directions in the face of ongoing environmental change. Phenotypic plasticity and the potential to modify morphology in response to changing environments is vital for all organisms. This includes marine fish, particularly in the face of climate change, and its plethora of effects on the marine habitat – from increasing sea temperatures, ocean acidification, and loss of food sources, to the loss of habitat and refuge this habitat provides.

1.3 Population genetics of fish in marine ecosystems

The dynamics and resilience of exploited marine species is important for their long-term survival. Application of genetics principles plays an important role in understanding factors underpinning the demography of fish populations. Traditionally, populations of marine species were believed to exhibit little genetic differentiation across large geographical distances (Ward, Woodward and Skibinski, 1994), due to the lack of physical barriers in marine seascapes, which combined with current flows, promote high levels of gene flow. In terms of evolution, many marine fish species were considered panmictic, characterised by long pelagic larval stages, and dispersal of young over large distances (Hemmer-Hansen *et al.*, 2014). Thus, marine populations were traditionally thought to be unaffected by stochastic/rapid genetic changes, with the primary evolutionary influences affecting population genetic structure being selection and gene flow, with modest evolutionary change over time. The potential and pressure for local adaptation was considered minimal, due to extensive dispersal rates of adults/larvae, and high levels of migration and large population sizes (Jennings, Kaiser and Reynolds, 2001).

Molecular genetics has provided novel approaches to understanding marine ecology at the genetic level of evolutionary and demographic dynamics of wild populations. Techniques using multiple

molecular markers reveal genetic structure at fine scales in marine fish populations (Ward, Woodward and Skibinski, 1994b), and a diversity of adaptive differentiation (Nielsen *et al.*, 2009).

Genetic studies have identified examples of genetic differentiation at local scales, with examples ranging from tens to hundreds of kilometres (Jorde *et al.*, 2003; Nielsen *et al.*, 2004; Olsen *et al.*, 2008). Compared to terrestrial and freshwater species, levels of genetic differentiation in marine fish are much smaller (Waples, 1998), but when considered in conjunction with environmental, temporal and spatial factors, become significant (Bekkevold *et al.*, 2005; Cimmaruta *et al.*, 2008).

Life history traits are important for determining population dynamics and directly impact the biomass of a species as well as demography. Adaptive responses in life history traits as a result of mortality induced by overfishing select for earlier age of maturity and increased reproduction (Trippel, 1995; Law, 2000; Olsen *et al.*, 2004b; Olsen *et al.*, 2004). Selection of these phenotypic traits may drive evolution of exploited fish stocks, with potential loss of genetic variation and selection of traits associated with low productivity (Hutchinson *et al.*, 2003; Olsen *et al.*, 2004b).

The importance of life history traits and the effects of exploitation are of interest in a marine conservation context due to the ecosystem-related effects that shifts in life history traits can have on competitive interactions, predator-prey dynamics and relative abundance of species (Law, 2000). Alterations in life history traits / phenotype may be caused by plastic (non-evolutionary) phenotypic responses, and adaptive (evolutionary) genetic responses, which may operate concurrently (Ernande, Dieckmann and Heino, 2004). The strength of the resulting genetic change influences the direction and reversibility of the shift in phenotype and how well a population can recover from exploitation (Law, 2000; M, S and Lennart, 2006; Walsh *et al.*, 2006).

It is difficult to separate plasticity of response, environmental effect and genetic changes, particularly in marine environments (Hendry, Farrugia and Kinnison, 2008; Naish and Hard, 2008). Linking evolution of genotype with phenotype is difficult because most phenotypes are coded for by a selection of genes that interact with one another and the environment. This challenge of separating environmental effects, phenotypic plasticity, and genotype is further compounded by the fact that many phenotypes are correlated and that selection for one phenotypic trait influences evolution of linked phenotypes (Naish and Hard, 2008).

1.4 Detecting population structure in marine populations

Marine fish populations have been commonly managed assuming complete panmixia of populations without giving consideration to life history traits and genetic differences of a population, and have been based on geographical area alone (Pampoulie *et al.*, 2011). The failure of the sustainable management of fish stocks was exemplified by the catastrophic depletion and collapse of Atlantic cod (Myers, Hutchings and Barrowman, 1996; Cook, Sinclair and Stefánsson, 1997) due to populations being viewed as singular units, rather than discrete biological units possessing subtle genetic differences, which are important for overall genetic health and resilience of the population (Stephenson and Kenchington, 2000). The main factors affecting gene flow and exchange of migrants between population subunits are geographical distance, current flow, and oceanic topography.

Studies using genetic markers to investigate population structure indicate the importance of considering genetic differentiation of populations, alongside life-history traits and ecological and environmental factors in order to best conserve marine fish populations and preserve biodiversity (Jorde *et al.*, 2003; Sarvas and Fevolden, 2005).

The fixation index statistic (F_{ST}) is used to measure levels of genetic differentiation within and between natural populations. It is calculated based on genetic polymorphisms arising from Single Nucleotide Polymorphisms (SNPs) or other genetic markers such as microsatellites. F_{ST} is an application of Wright's F statistics and is based on the differences between alleles at a given locus and the probability of an allele being inherited as "identical by descent" (or the probability that two individuals in a population are identical by descent). F_{ST} is an estimation of how closely related individuals in a population are to one another with values ranging between 1-0. A value of 0 represents complete panmixia, and a value of 1 indicates that the two populations share no genetic diversity, and all genetic variation is explained by the population structure.

In addition to using genetic data to estimate genetic differentiation between and within populations, population genetics and F_{ST} values can be used to estimate connectivity between putative populations and rates of migration and dispersal. An understanding of dispersal and interactions between individuals of a population is key to understanding the fundamental dynamics of ecological systems and gives rise to a myriad of population structure outcomes – from complete panmixia to distinct sub-populations (Rousset, 1997; Hanski, 1998). Dispersal of individuals is affected by a number of factors including population size, life-history traits (such as age at maturity), oceanographic variation, climate, and recruitment success (Hastings, 2004).

It is often difficult to interpret genetic data estimating connectivity between marine populations to demographic dispersal rates, mainly due to the fact that for many marine teleosts, F_{ST} genetic differentiation values are small and close to zero. Such low F_{ST} values are difficult to distinguish between moderate rates of dispersal and panmixia and low levels of genetic differentiation are likely due to high levels of dispersal, large population sizes and lack of physical barriers in the ocean. Population structure and connectivity has been demonstrated in numerous examples, with increasing numbers of studies based on high-throughput genetic sequencing data and providing evidence of fine scale population differentiation (Hess *et al.*, 2013; Benestan *et al.*, 2015; Rodríguez-Ezpeleta *et al.*, 2016).

1.4.1 Mitochondrial DNA (mtDNA) sequencing for phylogeographic analysis

Mitochondrial DNA (mtDNA) sequence data is a widely used tool in phylogeographic studies. This is due to its low/absent recombination rate, uniparental inheritance, conserved structure and relatively high evolutionary rate (Moritz, Dowling and Brown, 1987; Harrison, 1989). The analysis of within-species mtDNA variation can reveal information about the interconnectivity of populations and past demographic events such as population expansions and bottlenecks. The cytochrome oxidase I (CO1) gene is one of the most frequently employed mtDNA genes to investigate phylogeographic analysis and has been extensively used for evolutionary studies.

1.4.2 Restriction Site Associated DNA Sequencing (RAD-Seq) and related genotype by sequencing

The advent of Restriction Site Associated DNA Sequencing (RAD-Seq) and related genotype by sequencing (GBS) technologies have revolutionised the field of marine conservation genetics, enabling the study of genomes for which no prior genomic resources are available (Baird *et al.*, 2008; Davey *et al.*, 2011; Narum *et al.*, 2013). The technology works by sampling putative homologous regions of the genome from individuals in the population, then sequencing these subsets of the genome and identifying and genotyping Single Nucleotide Polymorphisms (SNPs). The technology has the ability to generate thousands of genome-wide SNP markers which can be used to study neutral population structure and local adaptation (Allendorf, Hohenlohe and Luikart, 2010). In the context of marine management, RAD-seq is a suitable tool for the study of previously unstudied, widely distributed non-model species, for minimal expense. It provides a rapid estimate of population differentiation, connectivity and adaptation and is becoming the “go-to” tool for informing conservation and

management decisions (Hess *et al.*, 2013; Larson *et al.*, 2014; Puebla, Bermingham and McMillan, 2014).

The large amount of SNP genetic markers generated by RAD-seq has advantages over older methods using microsatellites and allozymes, to detect genetic differentiation. This is because greater numbers of markers allow a more statistically robust investigation into genome wide variation. The thousands of SNPs generated overcome limitations of previous methods by improving accuracy of population genetic estimates, as well as gaining new insights into previously unexplored regions of the genome. Additionally, studies have shown that limitations imposed by the inability to detect weak genetic differentiation using older methods and microsatellite markers, have been overcome using RAD-seq (Benestan *et al.*, 2015).

Elucidating population genetic structure is a challenge in marine species due to the low levels of genetic differentiation that exist (usually F_{ST} is < 0.01). Studies looking at genetic structure of the American lobster (*Hommarus Americana*) have revealed negligible or non-existent levels of genetic differentiation among individuals in geographically separated regions, using microsatellites and allozymes (Kenchington *et al.*, 2009; Harding *et al.*, 2011; L. Tracey *et al.*, 2011). However, a more recent study of this species using RAD-seq found low, yet significant, genetic differentiation ($F_{ST} = 0.0020$) across SNPs in 11 populations, clustering according to two corresponding sampling locations. Genetic differentiation between populations at the two clusters ranged from $F_{ST} = 0.0010$ to 0.0037 (Benestan *et al.*, 2015). Whilst these values of genetic differentiation are low, they are biologically significant and may be associated with local adaptations (Palumbi, 1994; Nielsen *et al.*, 2004; Hess *et al.*, 2013; Benestan *et al.*, 2015; Rodríguez-Ezpeleta *et al.*, 2016). Although the oceans lack the characteristic physical barriers that lead to population differentiation in terrestrial and freshwater environments, there are other potential factors in the seascape that may give rise to the fine structuring that is apparent, such as ocean currents directing larval / adult dispersal or habitat shifts in response to changes in environment, predation risk and resource availability (Palumbi, 1994; Nielsen *et al.*, 2004; Lorenz and Carvalho, 2008).

It is therefore crucial when looking at population structure of marine fish, that low levels of genetic differentiation are acknowledged as being potentially significant in terms of reflecting changes over time in the natural environment. This is particularly important when considering differentiation signals that may indicate population divergence, which may occur over extremely long timeframes and operating across large geographical scales, in response to slowly changing environmental

conditions. Such changing environmental influences may not appear directly linked due to the complexity of underlying processes and a lag in “cause and effect” being apparent.

In summary, a genomic approach and analysis of population structure at a genetic level can provide insight regarding gene flow and genetic variation between populations of fish species at nearby locations reefs. Combined with morphological data, genomic analysis may detect genetic variation between fish populations, arising in response to environmental drivers. This research undertook this approach for two focal fish species in a study area as outlined in Sections 1.5 and 1.6.

1.5 The study area - the Chagos Archipelago Marine Protected Area

This research was conducted in the Chagos Archipelago Marine Protected Area (MPA), which is situated in the Central Indian Ocean (6° S, 71°30 E), part of the British Indian Ocean Territory (BIOT) (Fig 1.1). This remote coral reef ecosystem’s exceptional ecological value led to its designation as one of the world’s largest MPAs and it became a full no-take zone in 2010. Following the establishment of the MPA, all fishing in the BIOT waters became illegal, with the exception of recreational fishing in the waters surrounding Diego Garcia.

The MPA encompasses an area of approximately 640,000 km², comprising seven atolls and coral reefs covering an area of ~ 15,000 km². The Chagos Archipelago includes 25-50% of the total coral reef cover in the Indian Ocean and the largest contiguous intact reef area in the world (Sheppard *et al.*, 2012). The central feature of the Archipelago is the Great Chagos Bank, (the largest atoll on Earth), with the islanded atolls of Diego Garcia to the southeast, Salomon and Peros Banhos to the north and Egmont to the southwest, along with many submerged reefs. The only inhabited atoll is Diego Garcia, which is home to a US military base and personnel.

The Chagos Archipelago plays an important role in the biogeography of the Indian Ocean and acts as a ‘stepping stone’ location connecting marine populations between the western Indian Ocean and Indonesia (Obura, 2012). The remote location and lack of anthropogenic disturbance at these reefs makes them some of the last remaining, virtually ‘pristine’ coral reef environments in the world (Sheppard *et al.*, 2012).



Figure 1.1. Map showing the location of the Chagos Archipelago MPA (outlined in red) (Dunne *et al.*, 2014)

The remote location of the Chagos Archipelago coral reefs and lack of human interference has minimised some anthropogenic impacts (pollution, overfishing and industry) that have been detrimental to other global reef systems (Jackson *et al.*, 2001). Although the reefs were subjected to global warming events, they appeared to make a robust response following the global mass bleaching event of 1998 (Sheppard, 2003), 2015 and 2016 (Head *et al.*, 2019). Increasing temperatures in recent years gives cause for concern regarding the long-term resilience of the Chagos Archipelago coral reefs (Head *et al.*, 2019). However, currently, an abundance and diversity of reef fish (Graham *et al.*, 2013), low coral disease incidence and a lack of invasive, introduced marine species (Graham, Nash and Kool, 2011) and some of the cleanest water globally, indicates good reef health.

These factors all combine to create some of the world's most healthy coral reef and marine ecosystems worldwide. As such, the region provides an excellent study site to investigate baselines of ecosystem processes and marine populations. These attributes make the Archipelago an ideal location for understanding how species within reef communities might adapt to environmental change and degradation.

1.6 Investigating morphometric and genetic variation in two focal snapper species

The fish species selected for this study were large mesopredatory teleost lutjanid species: two spot red snapper (*Lutjanus bohar*) and humpback red snapper (*Lutjanus gibbus*). Both are widely distributed throughout the Indo-Pacific and exploited throughout their range. Lutjanids compete with other large piscivores, including reef sharks are primarily nocturnal feeders, and experience competition for the same food resources (I Nagelkerken *et al.*, 2000; Gillibrand, Harris and Mara, 2007).

Previous studies have shown that lutjanids have altered size and shape in response to environmental drivers, with this research mainly focussing on the effect of predation risk to the fish (Barley, Meekan and Meeuwig, 2017a; Barley, Meekan and Meeuwig, 2017b; Hammerschlag *et al.*, 2018). Lutjanids also exhibit behaviour associated with fear risk-effects exerted by shark presence / absence which affects their behaviour in terms of access to food, even if they are not directly predated by sharks (Creel and Christianson, 2008; Wirsing *et al.*, 2008).

The Lutjanidae are part of the Perciformes (perch-like) order of fish, comprising 17 genera and 109 species (Eschmeyer, Fricke and Van der Laan, 2016), largely distributed in tropical and sub-tropical regions. Worldwide, Lutjanidae are the most common predatory fish found on coral reefs (Allen, 1985). They have both ecological and economic importance, being harvested throughout their distribution by artisanal and commercial fisheries (Allen, 1985; Anderson and Allen, 2001; Marriott, Mapstone and Begg, 2007; Hollenbeck *et al.*, 2015). Lutjanids are nocturnally active predators, mainly feeding on other fish, crustaceans (including crabs, lobsters, stomatopods, and shrimps), molluscs (gastropods and cephalopods), and pelagic urochordates (Allen, 1985). In common with other marine tropical fish species, lutjanids have two life cycle phases, comprising a pelagic larval stage and reef-associated phase (Allen, 1985). They exhibit K-selected life history traits, being large, slow growing, and long-lived (Allen, 1985; Anderson and Allen, 2001), leaving them vulnerable to over-exploitation by fishing. They are gonochoristic, with adults reaching maturity at approximately 50% of maximum length. Batch spawning is nocturnal, and females spawn multiple times during a spawning season. Eggs and larvae are pelagic. Spawning occurs primarily between April and July, although some lutjanids spawn all year with peaks during spring and autumn (Grandcourt, Abdessalaam and Francis, 2006).

1.6.1 *Lutjanus bohar*

Lutjanus bohar (Forsskål, 1775) (two spotted red snapper), is a large piscivorous tropical snapper with wide-spread Indo-Pacific distribution. Adults have red/brown colouring (Figure 1.2) and juveniles have two characteristic ventral spots (Figure 1.3). It inhabits coral reefs and is often found on outer reef slopes and in deeper waters away from the reef slope, at depths between 10-70 m (Marriott, Mapstone and Begg, 2007).

Females mature at older age (9 years) and bigger size (420 mm Total Length (TL) on average) than males (1.36 years and 300 mm TL) (Marriott, Mapstone and Begg, 2007).

The species exhibits key criteria supporting its suitability as a candidate species to investigate the aims of this research project. It is a large fish, reaching a maximum length of 90 cm as an adult (Froese R and D., 2017) and occupies a high position in reef food webs, acting as both competitor and prey with reef sharks (Wetherbee, Crow and Lowe, 1997; Papastamatiou *et al.*, 2006, 2010).

L. bohar, in common with other lutjanids and reef sharks, feeds nocturnally (Nagelkerken *et al.*, 2000; Gillibrand, Harris and Mara, 2007), which is a common feeding pattern between competing mesopredators. Consequently, lutjanids experience stronger risk effects where reef shark populations are high, than would be apparent if they were diurnally active. Altered behaviour may also result without direct consumption when sharks are present (Creel and Christianson, 2008; Heithaus *et al.*, 2008). *L. bohar*'s wide-ranging diet allows the study species to shift diet selection, in response to increased or decreased risk effects caused by shark presence / absence.



Figure 1.2 Adult *Lutjanus bohar*
Photo-Sarah Speight: Flickr License



Figure 1.3 Juvenile *Lutjanus bohar*
Photo-Sally Pollack: FishWise

1.6.2 *Lutjanus gibbus*

Lutjanus gibbus (Forsskål, 1775) (the humpback red snapper), is a large carnivorous tropical snapper, with wide distribution throughout tropical Indo Pacific oceans (Figure 1.4). An active predator of other fish, *L. gibbus* also feeds on benthic invertebrates including lobsters, crabs, cephalopods, stomatopods, and echinoderms. Adults primarily inhabit coastal slopes on coral reefs, to depths of 150 m, forming large slow-moving/stationary aggregations during the day. Juveniles occupy seagrass beds and mixed sand and coral habitats of sheltered shallow reefs (Allen, 1985). *Lutjanus gibbus* grows to lengths of 50 cm, is red to grey in colouring, with a distinct caudal fin with rounded lobes. Juveniles have a characteristic large round black spot at the base of the caudal fin (Allen, 1985). Although information regarding spawning of *L. gibbus* is scant, it is thought that in common with other lutjanids, the spawning season occurs primarily during summer months, between May and October, with spawning events linked to lunar phases (Nanami *et al.*, 2010).



Figure 1.4 Adult *Lutjanus gibbus*

Photo – Ian Shaw: Reeflife survey

1.7 Aims

This research set out to investigate the genetic and morphometric variability of two focal species of lutjanid species in the Chagos Archipelago. The aims of the research are summarised below:

- 1) To use genetic variation at the cytochrome c oxidase subunit 1 (CO1) mitochondrial DNA (mtDNA) gene to investigate geographic variation across the Chagos Archipelago in the Indian Ocean, for *L. bohar*.
- 2) To assess morphological (size and shape) variation of the two focal species of *L. bohar* and *L. gibbus* between the Chagos atolls.
- 3) To undertake a genomic approach to generate single nucleotide polymorphism (SNP) markers to determine whether population structure between atolls in the Chagos exists.

Chapter 2 - Variability in the CO1 gene in populations of *L. bohar* in the Chagos Archipelago

2.1 Introduction

Molecular techniques such as DNA barcoding have routinely been used to identify species that may be difficult to differentiate based on morphology alone. The mitochondrial cytochrome oxidase 1 gene CO1 is a 658 bp long universal genetic marker used for DNA barcoding (Herbert *et al.* 2003). CO1 can be used to identify species across a broad range of phyla based on within sequence variation between species (Hebert, Ratnasingham and de Waard, 2003). Variation at the CO1 gene can be used for phylogenetic analysis to determine within species genetic variation. Phylogenetic analysis using CO1 molecular genetic information can also be used to provide information regarding genetic diversity within and between species, as well as inferring the historical and demographic connectivity of populations.

Snappers (fish belonging to the family Lutjanidae) can be morphologically difficult to identify, but CO1 DNA 'barcodes' have been successfully used to barcode all fish species (Ward, Hanner and Hebert, 2009), describe new species (Figueroa, 2008) and detect cryptic species (Smith *et al.*, 2008; Ward, Holmes and Yearsley, 2008), so provide a reliable means of identifying lutjanids. The two spotted snapper *L. bohar* is a large mesopredatory teleost, of commercial importance, that is widely distributed throughout the Indo Pacific where it inhabits tropical waters at depths of 10-180m (Froese R and D., 2017).

The aim of this research was to use genetic variation at the cytochrome c oxidase subunit 1 (CO1) mitochondrial DNA (mtDNA) gene to investigate geographic variation across the Chagos Archipelago in the Indian Ocean, for *L. bohar*. Based on previous research on marine fish, the hypothesis is that there will be minimal divergence between populations and the majority of individuals will share the same haplotype for the CO1 gene. A secondary aim was to confirm species identity for *L. bohar* samples that had been visually identified at sea.

2.2 Methods

2.2.1 Sample collection

Tissue samples were collected from *L. bohar* at four locations within the Chagos MPA. Samples were collected from the 2015 expedition to the Chagos Archipelago, carried out by the Centre for Marine Research at the University of Western Australia (UWA) at four atolls (Diego Garcia, Egmont, Salomon and Peros Banhos). Locations were distributed across a 200km² area of the Chagos Archipelago, with reefs being approximately 100-200 km apart (Figure 2.1).

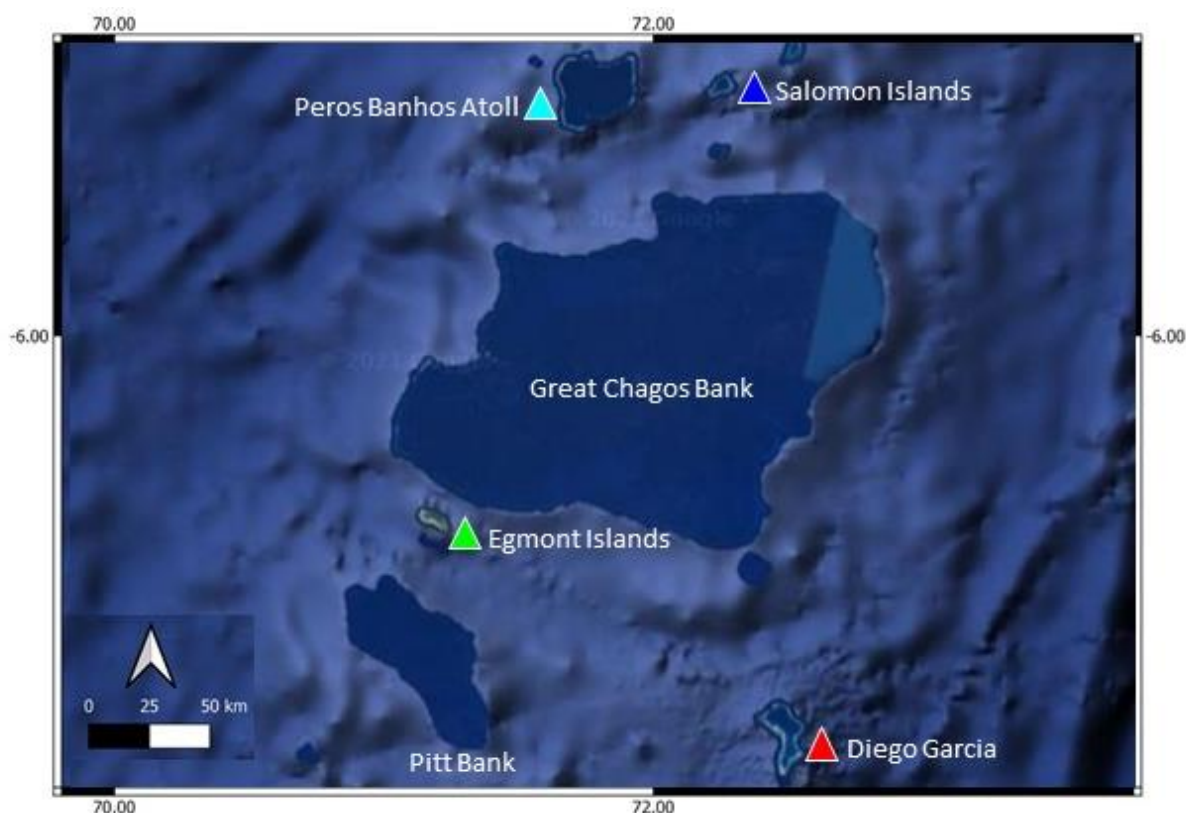


Figure 2.1 Sampling locations of *L. bohar* across the Chagos Archipelago, Indian Ocean. Sites are colour coded as follows: Diego Garcia – red, Egmont Islands– green, Peros Banhos atoll – turquoise, and Salomon Islands (blue).

Due to the protected status of the region, Foreign Commonwealth Office approval was secured for the research. Adult *L. bohar* were targeted and collected, using pole and line fishing methods by professional fishermen, following ethical approval from UWA and independent assessment by a licensed veterinarian. For individuals longer than 25 cm, a 0.5 cm³ punch biopsy of muscle tissue was taken from the ventral dorsal area of the fish and immediately stored in RNAlater™ (ThermoFisher Scientific). During sample excision, care was taken not to damage the fish or cause bleeding. Fish were subsequently tagged and released live. Samples were stored at 4°C and on arrival at the UK laboratory were transferred to - 80°C long-term storage, prior to processing for DNA extraction.

2.2.2 DNA extraction

Genomic DNA was extracted from 78 visually identified *L. bohar* dorsal muscle tissue samples using the Qiagen Dneasy Blood and Tissue Kit (Qiagen, UK). Genomic samples stored in RNAlater™ were thawed on ice, then dissected using sterile, disposable scalpels. The manufacturer's protocol was followed, with the addition of Riboshredder™ (Epicenter,US) following tissue lysis, to degrade unwanted RNA. Samples were eluted into 35 uL Qiagen AE elution buffer (10mM Tris-Cl, 0.5 mM EDTA, pH 9.0) giving an optimal DNA concentration of >20 ng/uL (determined using a Qubit® High Sensitivity assay and a Qubit® 3.0 fluorometer – Life Technologies).

To confirm sample quality, samples were run alongside a 1kb DNA Ladder (NEB Quickload®), on 1.2% agarose (Sigma) 0.5x TAE gel, using SYBR™ Safe (Invitrogen) DNA stain at 1:10,000 dilution. Following quality control of the genomic DNA, 78 *L. bohar* samples were submitted for CO1 sequencing.

2.2.3 CO1 sequencing

To confirm species identification, CO1 barcoding was carried out as described in (Kochzius *et al.*, 2010) using degenerate primers, based on universal CO1 primers for fish detailed in (Ward *et al.*, 2005): CO1-Fish-F (5'-TTC TCA ACT AAC CAY AAA GAY ATY GG-3') and CO1-Fish-R (5'- TAG ACT TCT GGG TGG CCR AAR AAY CA-3'). The degenerate primers amplified a 710 bp length of the CO1 sequence.

PCRs were carried out in a total volume of 15 uL, with each reaction comprising 7.5 uL (2X) Qiagen HotStarTaq® Master Mix, 0.5 uL CO1-Fish-F primer (10 uM), 0.5 uL CO1-Fish-R primer (10 uM), 4.5 uL deionised water, 2uL of DNA template. Thermal profile started with 95 °C denaturation step, followed by 30 cycles of 94 °C for 30 s, 59 °C for 30 s, 72 °C for 90 s, and final extension step of 72 °C for 7 mins. PCR reactions were confirmed by gel electrophoresis and amplification products purified using ExoSap-IT (Affymetrix) according to manufacturer's instructions.

Automated sequencing was carried out on the PCR products, in both forward and reverse directions, using an Applied Biosystems 3730xl DNA Analyser at the Department of Zoology, University of Oxford.

2.2.4 Sequence alignment

Sequences were paired and edited using chromatographic visualisations in Geneious (version 2021.1.1), (Biomatters Ltd, New Zealand). Samples with sequences that were of poor quality were discarded and not included in the multiple alignment step. Following manual editing of raw sequences, multiple alignments were compiled, using the Geneious alignment tool with default settings. Once

aligned, the 710 bp CO1 sequences were trimmed to a length of 660 bp. The Geneious Tree Builder tool was used to construct a circular phylogenetic tree, using neighbour-joining tree estimates and Tamura Nei genetic distances.

Sequences were “blasted” using the nucleotide sequence search tool at www.ncbi.nlm.nih.gov. All raw sequences were identified as *L. bohar*. Aligned sequences were downloaded as a FASTA format file for downstream haplotype network analysis.

2.2.5 Haplotype network construction

Haplotypes were extracted using the aligned FASTA format input file (created in Geneious) using R (4.0.3 version) and the **pegas** package. The aligned FASTA file, (including location and sample ID tags) was used to extract information about haplotypes and sequence variation, using the *haploFreq* and *nuc.div* functions. Genetic diversity within populations was estimated by calculating haplotype diversity (H) and nucleotide diversity (π). Haplotype diversity can be defined as the probability that two haplotypes drawn randomly from a population are not the same (Nei, 1987). Nucleotide diversity is defined as the mean number of nucleotide differences at each site in pairwise DNA comparisons (Nei, 1987). Relationships between haplotypes were estimated using the *haplotype* function in R to construct a median-joining haplotype network.

2.2.6 Neutrality tests

To assess whether mutations in the CO1 sequence are evolving neutrally (under mutation-drift equilibrium) or under non-neutral processes (directional selection or demographic expansion and reduction) Tajima’s D statistical test was used. This uses the frequency of segregating nucleotide sites and is based on the assumption that population expansion and reduction associated with non-neutral processes results in a shift in allele frequencies compared to that seen in neutral evolution. The analyses were implemented in R using the **pegas** package and the *tajima.test* function. P-values were generated using 1000 simulations under a model of selective neutrality.

2.3 Results

2.3.1 Sequence alignment

Sequencing was successful for 62 of the 78 *L. bohar* samples originally submitted for sequencing – poor quality sequences were discarded. Any sequence with missing bases was also removed at this stage, leaving 43 individuals for further analysis, distributed geographically as detailed in Table 2.1.

The sequence alignments for all 43 samples displayed a 99.8% pairwise identity, with none of the sequences varying by more than seven base pairs (Figure 2.2).

All CO1 sequences identified as originating from *L. bohar* when compared to published sequences on the NCBI Blast search facility, which confirmed the visual morphological identification of the specimens at time of sampling.

Table 2.1 Numbers of *L. bohar* used for CO1 sequence alignment at each location in the Chagos Archipelago

Sampling Location	<i>L. bohar</i> – number of individuals
Diego Garcia (DG)	16
Egmont (EG)	9
Peros Banhos (PB)	13
Salomon (SA)	5

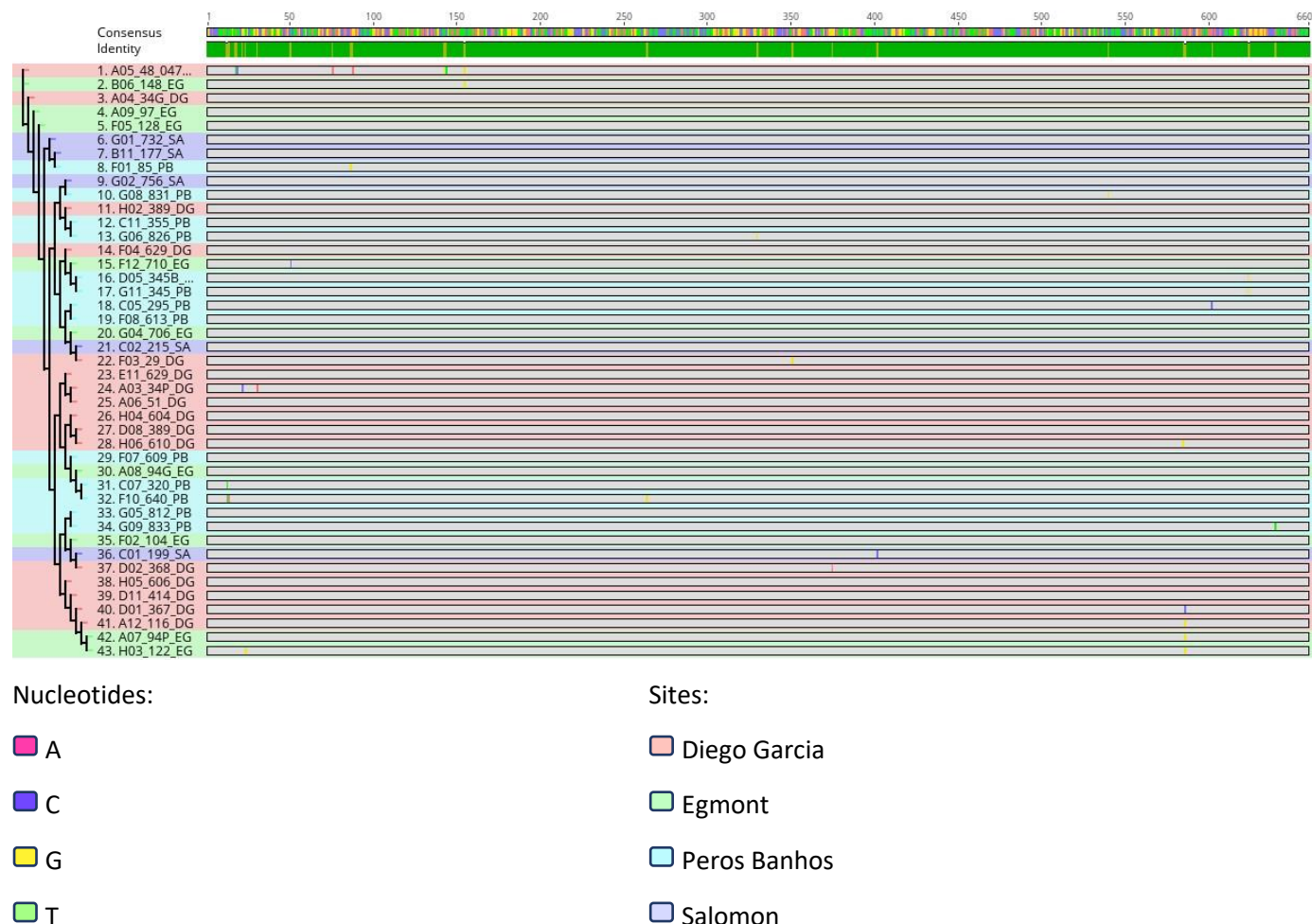


Figure 2.2 Multiple sequence alignments and phylogenetic tree belonging to the mitochondrial CO1 gene of *L. bohar*. Image highlights the nucleotide differences between sequences. Nucleotides are colour coded with four different colours. Sampling site for each sample is colour coded according to location.

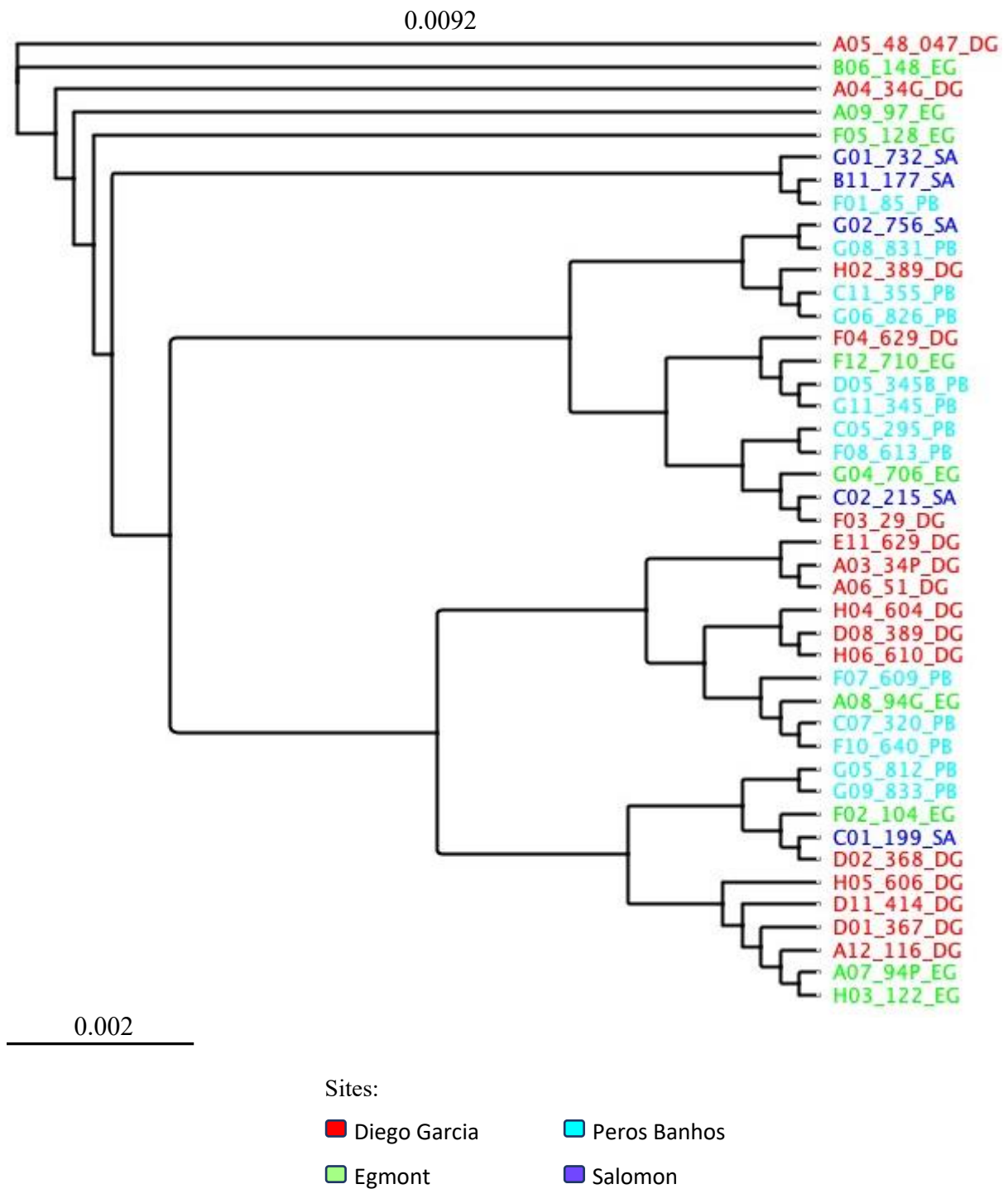


Figure 2.3 Phylogenetic neighbour-joining tree belonging to the mitochondrial CO1 gene of *L. bohar*. Tree was created using the Tamura Nei genetic distance model. The scale bar represents the genetic distance. Sample names are colour coded by sampling site.

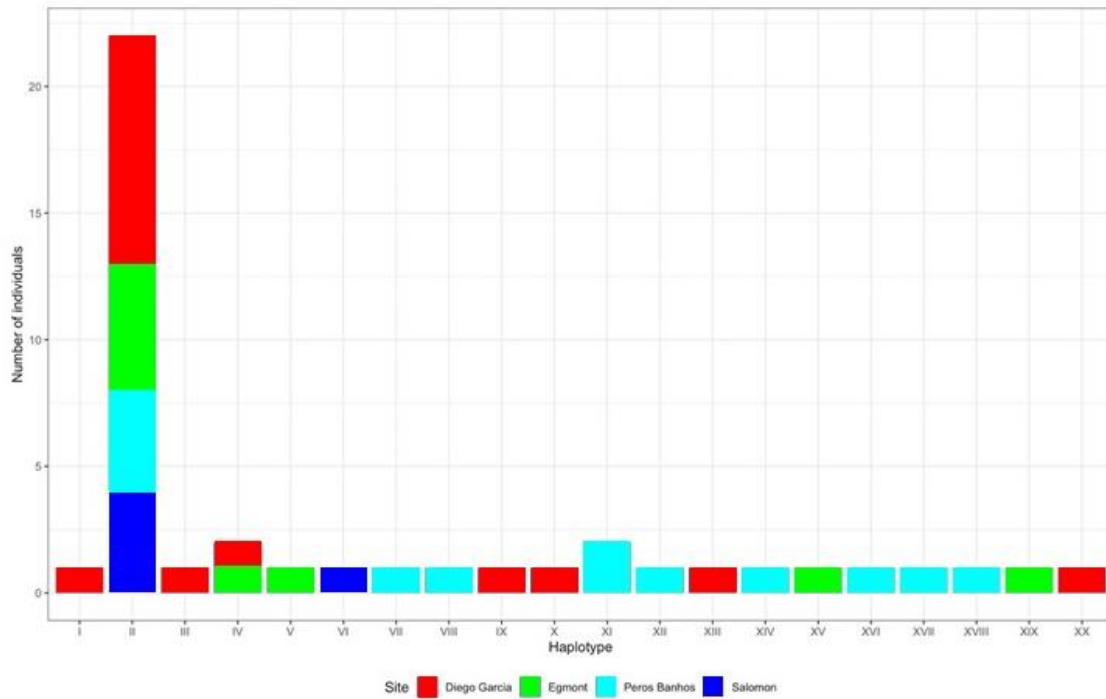


Figure 2.5 Histogram showing relative distribution of haplotypes for the CO1 gene in 43 individuals of *L. bohar* sampled at four locations across the Chagos MPA.

The number of haplotypes for each location and estimates of haplotype diversity are shown in Table 2.2. All populations contained at least 2 haplotypes and the most abundant haplotype (haplotype II) was found in 51.2 % of individuals. The second most abundant haplotypes (IV and XI) were each found in 4.7% of individuals.

Table 2.2 Genetic variation within sites based on a 660 bp sequence of the mitochondrial CO1 gene for *L. bohar*. n sample size, n_hap number of haplotypes, H haplotype diversity, π nucleotide diversity. D – Tajima’s test for neutrality (all p values for D were not significant).

Site	n_	n_hap	H	H variance	π	D
Diego Garcia	16	8	0.7	0.016	0.0026	-2.142
Egmont	9	5	0.7222	0.024	0.0015	-1.149
Peros Banhos	13	9	0.9273	0.003	0.0026	-1.493
Salomon	5	2	0.4	0.0570	0.0006	-0.816

Sample sizes per location are unequal but the proportional number of haplotypes were comparable for three sites (Diego Garcia, Egmont, and Salomon), when different sample sizes are accounted for.

The exception was Peros Banhos which has the highest proportion of haplotypes when considering sample size.

For the entire region, haplotype diversity was relatively high ($H = 0.74$) and nucleotide diversity was low ($\pi = 0.0021$), with these values being similar to other studies conducted on lutjanid species (Gaither *et al.*, 2009). Haplotype diversity was greatest in Peros Banhos ($H = 0.93$), whilst Salomon had much lower genetic diversity with $H = 0.4$ (this may be due to the very small sample size of 5 for Salomon). Nucleotide diversity was low for all sites. Diego Garcia, Egmont and Peros Banhos had nucleotide diversity values ranging from 0.0016 to 0.0026. Salomon had the lowest nucleotide diversity ($\pi = 0.0006$), again, possibly due to small sample size.

2.3.3 Neutrality tests

High haplotype diversity and low nucleotide diversity values can be suggestive of a bottleneck followed by population expansion, but this was not supported by the Tajima's D test for neutrality of mutations of the CO1 region.

The neutrality test for individuals at each location is shown in Table 2.2. Tajima's D value was not statistically significant for any location with all P values being greater than 0.05. This indicates no evidence of recent bottlenecks or occurrences of selection, and that polymorphisms are selectively neutral for the region.

2.4 Discussion

Our results showed relatively little differentiation of *L. bohar* across the four locations sampled in the Chagos Archipelago. The molecular diversity indices (haplotype and nucleotide diversity) are similar between most populations. Haplotype diversity was relatively high for three of the sites (Diego Garcia, Egmont and Peros Banhos), ranging from 0.7 to 0.927, which is supported by similar H values for other lutjanid species (Garber *et al* 2004, Gomes *et al*, 2008 and Zhang *et al*, 2006). The low haplotype diversity value at Salomon is generated from a small sample size, so should be treated with caution as if greater numbers of individuals had been considered, a different haplotype profile may have been identified. Although haplotype diversity tends to be relatively high for most sites in the region, the nucleotide diversity indices are low, indicating small differences between haplotypes, and these values are also supported by results in other studies on lutjanid species.

Whilst the most common haplotype (haplotype II) is shared across all four sites, rarer haplotypes are attributable to single nucleotide differences and are often singletons. The distribution and prevalence of haplotypes indicates that geographical distance between sites does not appear to be a factor in determining which haplotypes occur at each location. High haplotype diversity and sharing of the common haplotype indicates good connectivity between sites.

The lack of genetic variation and sharing of haplotypes for the CO1 gene across the Chagos atolls indicates that *L. bohar* is genetically homogeneous throughout the geographic region and that there is good connectivity between populations at different sites. These findings are not surprising, given that gene flow and mixing of individuals between populations at different atolls is primarily mediated via larval dispersal, which potentially enables long-distance mixing of marine organisms (Underwood, Travers and Gilmour, 2012). The atolls of Chagos studied in this research are separated by distances of approximately 150-200 km. The larvae of lutjanids have pelagic larval durations that enable the dispersal of individuals over such distances (Harrison *et al.*, 2012; Green *et al.*, 2015, Muths *et al.*, 2012, Herrón and Zapata, 2002). Therefore, assuming that there are no other barriers to dispersal, it is likely that individuals can move freely across the entire region.

Connectivity between atolls may also (to a lesser degree) be mediated via adult movement, and there is evidence that some lutjanid species migrate large distances, (tens to hundreds of km), most likely influenced by the requirement to find prey and gathering for spawning events (Green *et al.*, 2015). It is therefore possible that occasional adult migration between atolls in addition to pelagic larval dispersal between sites, is an additional means by which genetic homogeneity is maintained across the overall Chagos site.

In conclusion, the findings support the hypothesis that there is minimal divergence between populations, with the majority of individuals sharing the same haplotype for the CO1 gene.

Chapter 3 - A morphometric analysis of *L. bohar* and *L. gibbus* within the Chagos Archipelago

3.1 Introduction

The evolution of the anatomical structure of any organism is driven by the complex interactions between environmental conditions, genetic traits and phenotypic plasticity (Lande, 2009). Phenotypic plasticity describes the potential of a genotype to generate a variety of different phenotypes in response to different environmental drivers. These phenotypic changes can be morphological, behavioural, physiological, developmental, or biochemical. In the context of evolution and species survival, phenotypic plasticity is important because it allows an organism to acclimatise or adapt within a single generation, in response to environmental cues within a changing and heterogeneous environment, such that the resulting phenotype is optimal for the altered environment (Pigliucci, 2005).

Fish are a highly plastic group of organisms and have a great ability to modify their phenotype in response to environmental drivers, resulting in morphological variation (Olsson and Eklöv, 2005; Eklöv and Svanbäck, 2006; Chapman, 2015). Multiple abiotic and biotic environmental factors influence fish phenotype, and these include temperature (Volkoff and Rønnestad, 2020) light (Kröger and Fernald, 1994), habitat complexity (Garduno-Paz, Couderc and Adams, 2010), food availability (Abaad *et al.*, 2016), predation risk (Hammerschlag *et al.*, 2018), current flow (Collin and Fumagalli, 2011) and prey biology (Davis, Pusey and Pearson, 2012).

Understanding the degree of morphological variation within marine fish populations is important, in order to understand and predict the impacts of anthropogenic disturbances. The ability of marine species to acclimatise to increasingly severe and frequent climate events and human impacts is an area of research that is key to ensuring the sustainability of global fish populations. To understand the effects of changing environments, it is first necessary to establish current norms of spatial morphological variation for marine species.

In experiments conducted at a natural scale, it can be difficult to isolate and control for multiple environmental variables that give rise to morphological variation. Specifically, most marine environments are subject to multiple anthropogenic disturbances such as over-fishing, pollution, and habitat degradation. To overcome this issue and to minimise these direct anthropogenic influences,

studies can be carried out at one of the remaining reef locations that are still considered as near to pristine as possible in the current global climate.

The Chagos Archipelago is situated in the middle of the Indian Ocean and its remote location and status as a Marine Protected Area, makes it an excellent reference site for conducting studies isolated from direct anthropogenic influences (Sheppard *et al.*, 2012).

This study investigated morphological variation (size and shape) for two focal snapper species (*Lutjanus bohar* and *Lutjanus gibbus*) at four atolls within the Chagos Archipelago (Diego Garcia, Egmont, Salomon and Peros Banhos). Both lutjanid species are exploited throughout their range (Marriott, Mapstone and Begg, 2007) and given their high position in trophic cascades (Froese R and D., 2017) and the impact they have on energy transfer in coral reef ecosystems, understanding the potential effects of environmental change is essential. Lutjanid species (including the two focal species) have previously been shown to acquire altered phenotypes in response to altered predation risk at other reef locations in the Indian Ocean (Barley, Meekan and Meeuwig, 2017a; Barley, Meekan and Meeuwig, 2017b; Hammerschlag *et al.*, 2018). However, baseline measures of morphological variation for other environmental drivers have not been recorded.

The aim of this research project was to undertake a comparative morphological study on two focal snapper species (*Lutjanus bohar* and *Lutjanus gibbus*) at atolls within the Chagos Archipelago to determine if any significant variation in morphology (size and shape) could be detected. The null hypothesis is that there is no significant difference in size or shape for either *L. bohar* or *L. gibbus* at different sites across the Chagos Archipelago.

3.2 Methods

3.2.1 Sample collection

Samples of *L. bohar* and *L. gibbus* were collected during two expeditions to the Chagos Archipelago, carried out by the Centre for Marine Research at the University of Western Australia (UWA) in November 2015 and May 2016 (Figure 3.1). Due to the protected status of the region, Foreign Commonwealth Office approval was secured for the research. Samples were collected from four atolls: Diego Garcia, Egmont, Salomon Islands and Peros Banhos. Key metrics including weight, fork length (body length from mouth to fork), body depth and circumference, and Go-Pro videos and

photos, were recorded at the time of capture. It was not always possible to collect comprehensive information for every specimen. Once sample data had been collected, live fish were then returned to the sea.

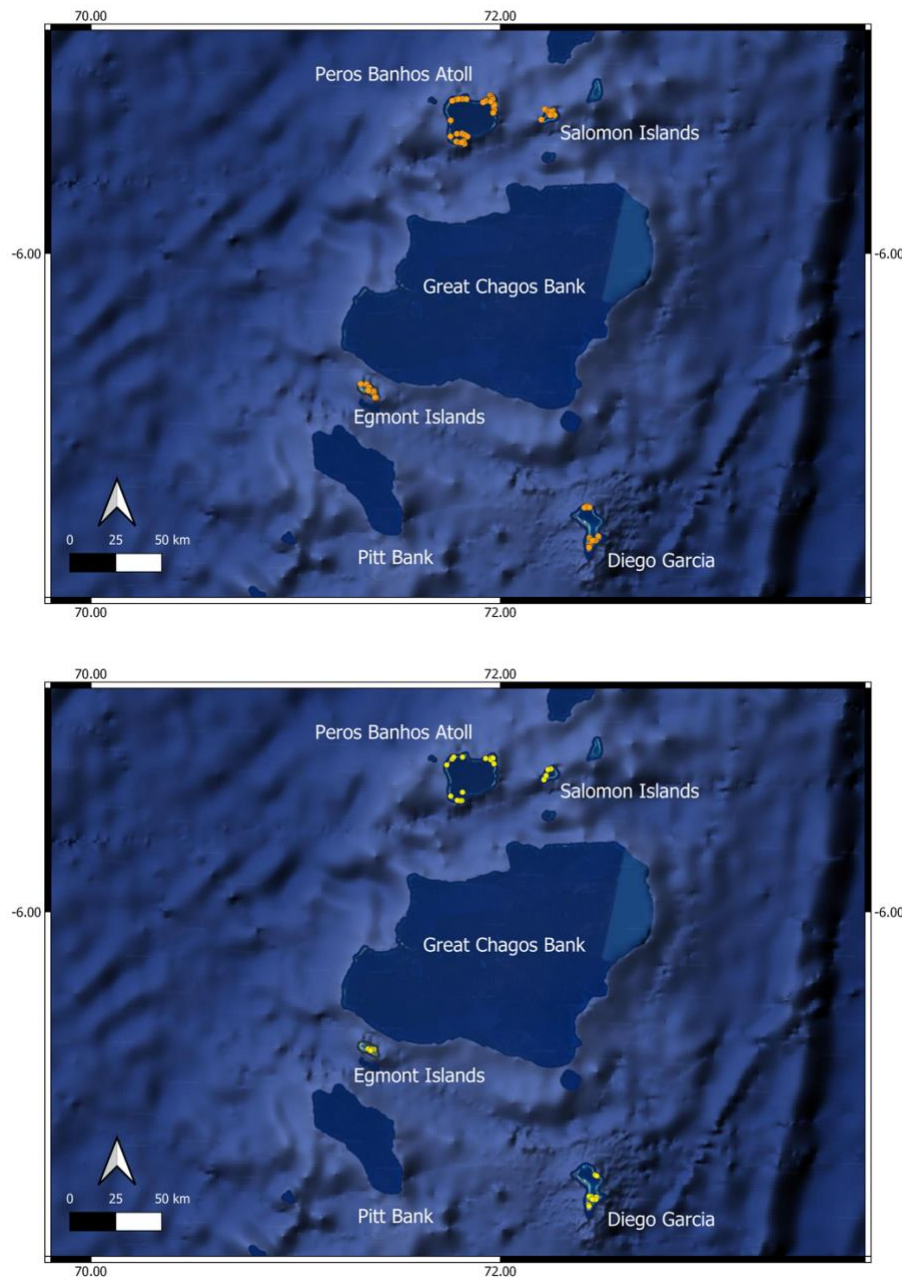


Figure 3.1 Maps showing the sampling locations in the Chagos MPA for *L. bohar* (top) (denoted by orange dots) and *L. gibbus* (bottom) (denoted by yellow dots) collected during expeditions in 2015/16

Raw data was checked for erroneous recordings that were attributable to sample recorder error. Based on existing morphological data available for both species (Froese R and D., 2017) any samples exceeding maximum fork lengths of 90 cm for *L. bohar* and 50 cm for *L. gibbus* were removed. Errors in the recording of body depth were also detected for some samples (which could be traced back to specific recorders). For these samples, body width had been mistakenly measured instead of body

depth. The incorrectly recorded raw data for body depth in these cases meant that individuals would appear to be less deep bodied for a given fork length than they were. All individuals with a relative ratio of fork length to body depth exceeding the published maximum of 3.5 (which also broadly equated to the ratio of fork length to body depth in the correctly measured samples in our data set) were re-measured using images taken from the video / photo footage recorded at time of sampling. Fork length was recalculated using a conversion for the onscreen measurements in Excel. Any individual with an erroneous measurement, that did not have an image recorded, was discarded from the dataset. In total, images were used to re-calculate measurements for 45 *L. bohar* and 13 *L. gibbus* samples.

Once erroneous measurements had been re-calculated, scatterplots were generated in R to detect any further outliers, and the corresponding samples were subsequently removed from the data for downstream analysis. Geographical distributions of the samples for each species included in the morphometric analysis are shown in Table 3.1.

Table 3.1 Numbers of samples (n) collected at each sampling location and used for morphometric analysis

Sampling Location	<i>L. bohar</i> (n)	<i>L. gibbus</i> (n)
Diego Garcia (DG)	32	43
Egmont (EG)	22	23
Peros Banhos (PB)	63	27
Salomon (SA)	12	8
Total individuals sampled	129	101

3.2.2 Size analysis

All data analysis was carried out for both species using R version 3.5.2 (R Core Team (2021)). Summary statistics were generated to calculate the mean, minimum and maximum values for weight, circumference, fork length and body depth at each sampling location. It should be noted that not all samples had data recorded for every metric, so sample numbers at each site were not always constant.

Histogram plots were generated to visualise the distribution of the data and the raw data was then log10 transformed to normalise the data. Using the **rcorr** function in R, a correlation matrix using Pearson's correlation coefficients, was created to test for the strength of correlation between size variables.

To test for differences in size between sampling sites, using the **aov** function in R, an ANOVA (Analysis of Variance) was carried out for each independent size variable (weight, circumference, body depth and fork length), using the location as the predictor variable. Variables for individuals that were missing data were automatically omitted from the ANOVA calculation. Box plots were generated using the log10 transformed data to visualise variations in the means for each size variable at different sampling locations. Where the ANOVA results indicated a statistically significant difference for a size metric between sites ($p < 0.05$), a Tukey test was carried out using the **emmeans** package in R to determine which populations differed from each other.

3.2.3 Shape analysis

An Analysis of Covariance (ANCOVA) test was carried out in R to test for differences in shape between sample sites. The log10 transformed data was used for this analysis. Sampling location was used as the predictor variable and length was used as the covariable. The remaining three response variables of weight, body depth and circumference were independently regressed against length and location in separate ANCOVA analyses. For each response variable that indicated a significant difference among sites ($p < 0.05$), a *post hoc* Tukey test was used to determine which populations differed from one another.

To eliminate potential issues associated with collinearity (which occurs when variables are highly correlated with one another) a residual analysis model was run. Residuals were extracted from the regression of each response variable on fork length and an ANOVA conducted on the residuals.

3.3 Results - size variation

3.3.1 Summary size statistics for *L. bohar* and *L. gibbus*

Mean values for all metrics indicate that *L. bohar* at Egmont are bigger than at all other atolls, with smallest individuals found at Diego Garcia (Table 3.2). There is an exception to this trend when considering the weight variable, with mean values being smallest at Peros Banhos than all other atolls; however, this difference in size is marginal and equates to a 30g difference in mean weight between Diego Garcia and Peros Banhos.

Table 3.2 Mean + Standard Error and minimum and maximum values for weight (kg), fork length (cm), circumference (cm) and body depth (cm) for *L. bohar* at each sampling location. N is the sample size at each site.

Location	Metric	N	Mean	SE	Min	Max
Diego Garcia	weight (kg)	21	1.62	0.33	0.3	7
Egmont		19	2.53	0.29	0.4	5.5
Peros Banhos		63	1.59	0.13	0.2	5.1
Salomon		12	2.34	0.40	1	5.2
Diego Garcia	fork length (cm)	32	35.09	2.02	14.5	70
Egmont		20	47.95	2.24	23	65
Peros Banhos		64	41.21	1.13	18	62
Salomon		12	46.67	2.82	31	67
Diego Garcia	circumference (cm)	32	25.98	1.46	11	51
Egmont		21	35.64	1.48	18	45
Peros Banhos		63	29.98	0.83	14	49.5
Salomon		12	33.33	1.71	24	46.5
Diego Garcia	body depth (cm)	32	11.24	0.62	5	21.39
Egmont		21	15.23	0.63	8	20
Peros Banhos		64	13.30	0.38	6	24.8
Salomon		12	14.92	0.84	11	22

For all variables measured for *L. gibbus*, mean values also showed that the smallest individuals are at Diego Garcia. For *L. gibbus*, the biggest fish are found at Salomon (Table 3.3). However, the sample size for *L. gibbus* at Salomon is small ($n = 8$), so it is possible that these results are skewed by this. It should be noted that individuals at Peros Banhos (the atoll adjacent to Salomon) are almost as big as individuals at Salomon, when looking at mean size metrics. The sample size at Peros Banhos ($n=26$) is much greater than that at Salomon, so the data has less variation for each size metric.

Table 3.3 Mean + Standard Error and minimum and maximum values for weight (kg), fork length (cm), circumference (cm) and body depth (cm) for *L. gibbus* at each sampling location. N is the sample size at each site.

Location	Metric	N	Mean	SE	Min	Max
Diego Garcia	weight/ kg	34	0.70	0.05	0.30	1.30
Egmont		21	0.89	0.10	0.30	2.20
Peros Banhos		24	0.88	0.06	0.40	1.70
Salomon		8	0.91	0.09	0.50	1.20
Diego Garcia	fork length/cm	43	24.47	0.84	16.5	41.0
Egmont		22	28.32	1.12	19.0	40.0
Peros Banhos		26	30.96	0.65	24.0	36.0
Salomon		8	33.38	1.54	28.0	38.5
Diego Garcia	circumference / cm	42	19.4	0.64	13.0	28.0
Egmont		22	21.7	0.81	16.0	29.0
Peros Banhos		25	25.4	0.69	20.0	33.0
Salomon		8	25.81	0.93	22.0	28.5
Diego Garcia	body depth / cm	41	7.37	0.48	2.0	13.0
Egmont		22	9.95	0.54	3.0	15.0
Peros Banhos		26	10.65	0.57	2.0	15.0
Salomon		8	12.25	0.31	11.0	13.0

3.3.2 ANOVA model to test for differences in size between sites for *L. bohar*

The ANOVA model for *L. bohar* data showed that there were statistically significant differences in weight, circumference, body depth and fork length, in relation to sampling location (see Table 3.4).

The results showed that there were statistically significant differences between location and size metrics as follows: mean weight ($F(3,111) = 4.339$, $p = 0.0062$); mean circumference ($F(3,124) = 9.112$, $p = <0.0001$); mean body depth ($F(3,125) = 9.161$, $p = <0.0001$) and fork length ($F(3,124) = <0.0001$). The most significant differences were for variations in mean fork length between sites ($p = <0.0001$), circumference ($p = <0.0001$) and body depth ($p = <0.0001$). The effect of location on weight ($p = 0.0062$) was less statistically significant compared to other size variables.

Table 3.4 ANOVA comparing mean size metrics between sampling locations for *L. bohar*.

P value significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Size variable	Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	Location	3	1.097	0.3658	4.339	0.0062 **
	Residuals	111	9.358	0.0843		
Circumference	Location	3	0.3173	0.1058	9.112	1.7e-05 ***
	Residuals	124	1.4392	0.0116		
Body depth	Location	3	0.3244	0.1081	9.161	1.59e-05 ***
	Residuals	125	1.4754	0.0118		
Fork length	Location	3	0.3273	0.1092	8.606	3.11e-05 ***
	Residuals	124	1.5720	0.0127		

3.3.3 Post hoc Tukey tests to test for differences in size of *L. bohar* between sites

The full results for Tukey's HSD Test for multiple comparisons are listed in Appendix Chapter 3 (Tables A3.1 to A3.4), with the significant differences summarised in the box plots below.

The mean value of *L. bohar* weight was significantly different between Egmont and Diego Garcia ($p = 0.035$), and between Egmont and Peros Banhos ($p = 0.027$), (Table A3.1, Figure 3.2). This reflects the fact that Egmont had the heaviest mean *L. bohar* with Diego Garcia and Peros Banhos having the lightest.

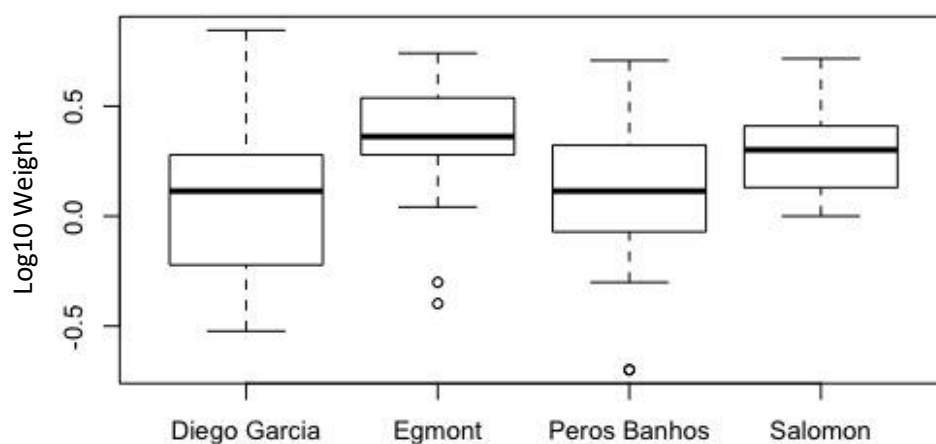


Figure 3.2 Box plot showing average log weight differs between sites for *L. bohar*. ANOVA on weight yielded significant variation between sites ($p=0.006$). Post hoc Tukey HSD indicates significant differences for weight between Diego Garcia vs Egmont ($p = 0.035$, 95% C.I. = 0.0130, 0.4926) and Peros Banhos vs Egmont ($p= 0.027$, 95% C.I. = -0.4137, -0.0173).

Tukey's HSD Test for multiple comparisons found that the mean circumference of *L. bohar* was significantly different between Diego Garcia and all locations - Diego Garcia and Egmont ($p = <0.0001$), Diego Garcia and Peros Banhos ($p = 0.0154$), Diego Garcia and Salomon ($p = 0.0060$) as well as between Peros Banhos and Egmont ($p = 0.0265$) (Table A3.2, Figure 3.3).

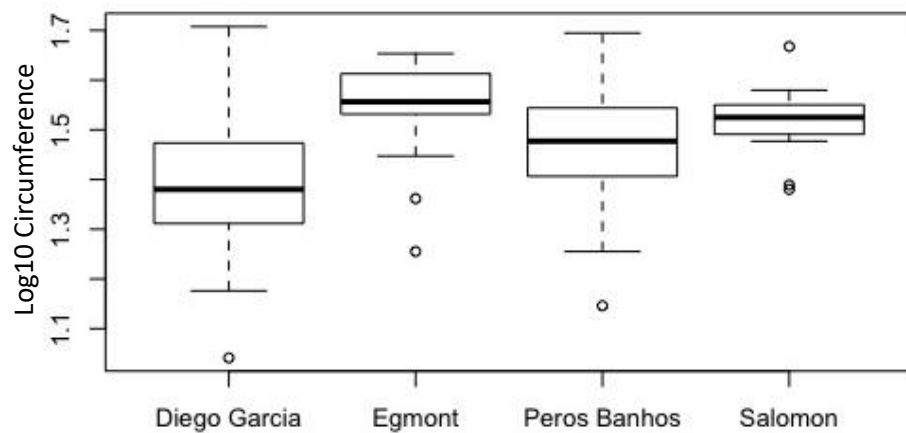


Figure 3.3 Box plot showing how average log circumference differs between sites for *L. bohar*. ANOVA on circumference yielded significant variation between sites ($p = <0.0001$). Tukey HSD indicates significant differences for circumference between Diego Garcia vs Egmont ($p = <0.0001$, 95% C.I. = 0.0694, 0.2269) and Diego Garcia vs Peros Banhos ($p = 0.0154$, 95% C.I. = 0.0101, 0.1318), Diego Garcia vs Salomon ($p = 0.0060$, 95% C.I. = 0.0268, 0.2168) and Peros Banhos vs Egmont ($p = 0.0265$, 95% C.I. = -0.1479, -0.0065).

Tukey's HSD Test for multiple comparisons found that the mean value of body depth for *L. bohar* was significantly different between Diego Garcia and all locations – Diego Garcia and Egmont ($p = <0.0001$), between Diego Garcia and Peros Banhos ($p = 0.0041$) and between Diego Garcia and Salomon ($p = 0.0019$). There was no significant difference between any other locations (Table A3.2, Figure 3.4).

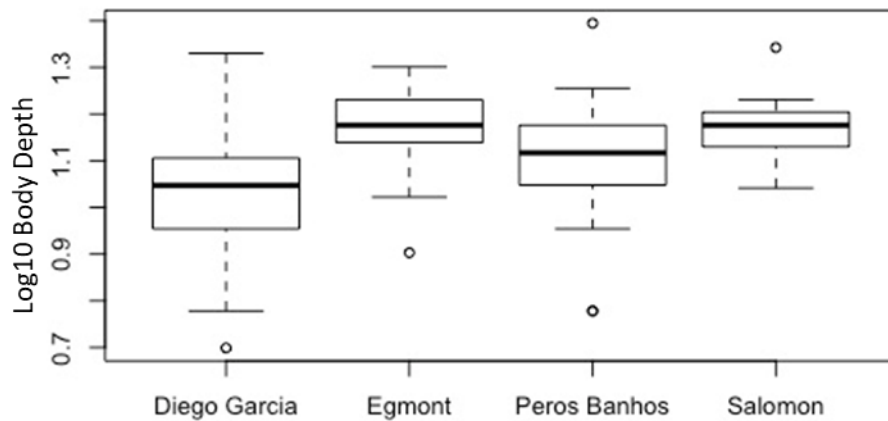


Figure 3.4 Box plot showing how average log body depth differs between sites for *L. bohar*. ANOVA on body depth yielded significant variation between sites ($p = <0.0001$). Tukey HSD indicates significant differences for body depth between Diego Garcia vs Egmont ($p = <0.0001$, 95% C.I. 0.0638, 0.2227), between Diego Garcia vs Peros Banhos ($p = 0.0041$, 95% C. I. 0.0200, 0.1425) and between Diego Garcia vs Salomon ($p = 0.0019$, 95% C.I. 0.1357, 0.0399).

Tukey's HSD Test for multiple comparisons found that the mean value of fork length for *L. bohar* was significantly different between Diego Garcia and all locations. Diego Garcia and Egmont ($p = <0.0001$), Diego Garcia and Peros Banhos ($p = 0.0074$) and between Diego Garcia and Salomon ($p = 0.0028$). No other locations had significant differences in mean fork length (Table A3.4, Figure 3.5).

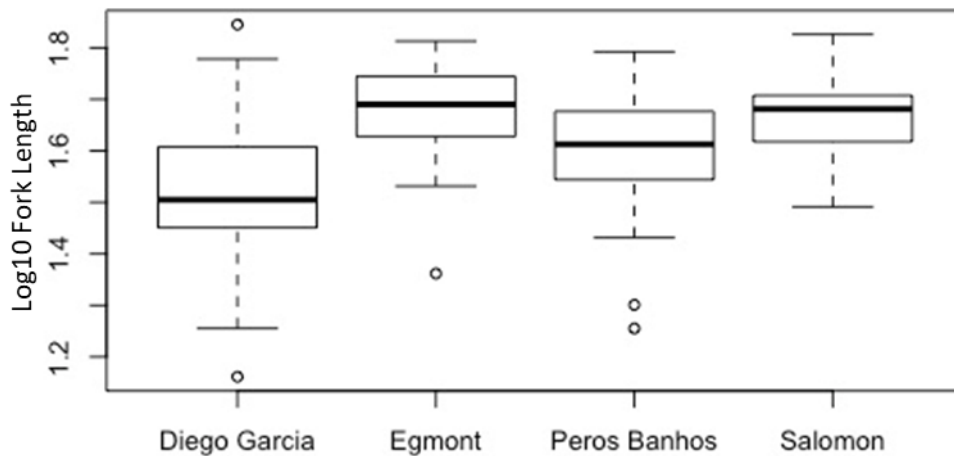


Figure 3.5 Box plot showing how average log fork length differs between sites for *L. bohar*. ANOVA on fork length yielded significant variation between sites ($p = <0.0001$). Tukey HSD indicates significant differences for fork length Diego Garcia vs Egmont ($p = <0.0001$, 95% C.I. 0.0625, 0.2297), between Diego Garcia vs Peros Banhos ($p = 0.0074$, 95% C.I. 0.0163, 0.1433) and between Diego Garcia vs Salomon ($p = 0.0028$, 95% C.I. 0.0368, 0.2353).

3.3.4 ANOVA model to test for differences in size between sites for *L. gibbus*

The ANOVA model for *L. gibbus* data showed that there were statistically significant differences in weight, circumference, body depth and fork length, in relation to sampling location (see Table 3.5).

The results showed that there was a statistically significant differences between location and size metrics as follows: mean weight ($F(3, 86) = 3.035$, $p = 0.0335$); mean circumference ($F(3,96) = 15.72$, $p = <0.0001$); mean body depth ($F(3,96) = 8.784$, $p = <0.0001$) and mean fork length ($F(3,98) = 15.41$, $p = <0.0001$).

Table 3.5 ANOVA comparing mean size metrics between sampling locations for *L. gibbus*.

P value significance codes represented by: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Size variable	Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	Location	3	0.2593	0.0864	3.035	0.0335 *
	Residuals	86	2.4497	0.0285		
Circumference	Location	3	0.2921	0.0974	15.72	2.16e-08 ***
	Residuals	96	0.5948	0.0062		
Body depth	Location	3	0.940	0.3133	8.784	3.33e-05 ***
	Residuals	96	3.424	0.0357		
Fork length	Location	3	0.285	0.0950	15.41	2.74e-08 ***
	Residuals	98	0.604	0.0062		

3.3.5 Post hoc Tukey tests to test for differences in size of *L. gibbus* between sites

The full results for Tukey's HSD Test for multiple comparisons are listed in Appendix Chapter 3 (Tables A3.5 to A3.8), with the significant differences summarised in the box plots below.

Tukey's HSD Test for multiple comparisons found that the mean value of weight was not significantly different between sites (Table A3.5, Figure 3.6)

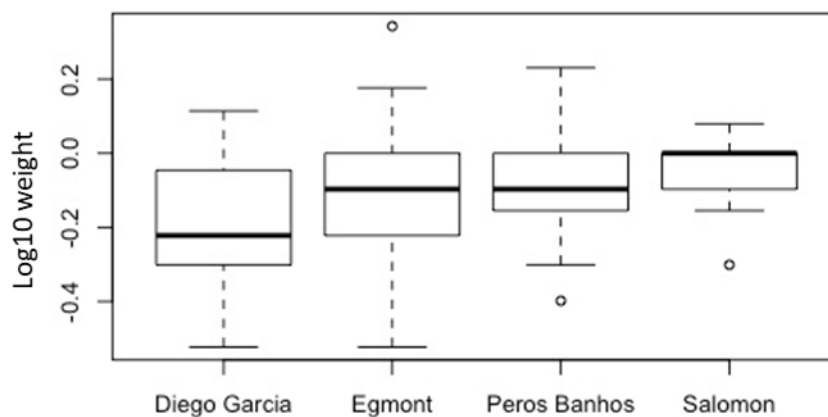


Figure 3.6 Box plot showing how average log weight differs between sites for *L. gibbus*. ANOVA on fork length yielded significant variation between sites ($p = 0.0335$). However, Tukey HSD indicates that weight is not significantly different between any individual sites ($p > 0.05$ for all pairwise site comparisons).

Tukey's HSD Test for multiple comparisons found that the mean value of circumference was significantly different between Diego Garcia and all locations. Diego Garcia and Egmont ($p = 0.0419$, 95% C.I. = 0.0014, 0.1081), Diego Garcia and Peros Banhos ($p = <0.0001$, 95% C.I. = 0.0689, 0.1717), Diego Garcia and Salomon ($p = <0.0001$, 95% C.I. = 0.0570, 0.2082) and Peros Banhos and Egmont ($p = 0.0231$, 95% C.I. = 0.0066, 0.1244), (Table A3.6, Figure 3.7).

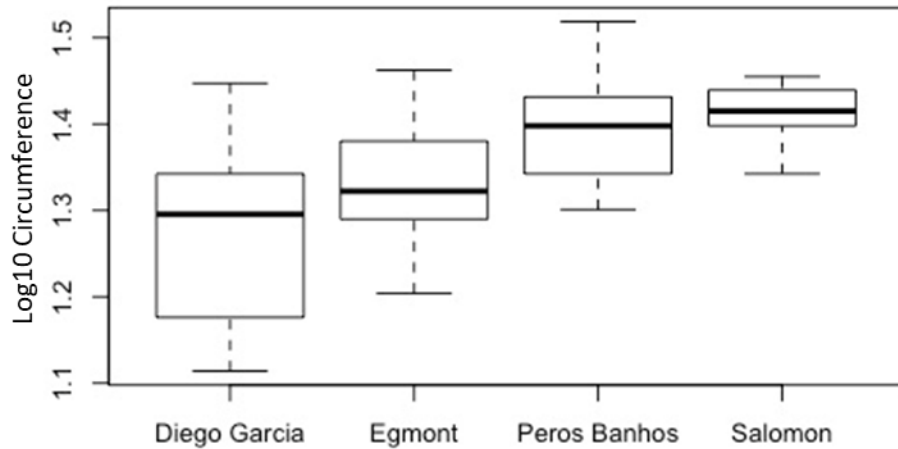


Figure 3.7 Box plot showing how average log circumference differs between sites for *L. gibbus*. ANOVA on circumference yielded significant variation between sites ($p = <0.0001$). Tukey HSD indicates significant differences for circumference between Diego Garcia and Egmont ($p = 0.0419$, 95% C.I. = 0.0014, 0.1081), Diego Garcia and Peros Banhos ($p = <0.0001$, 95% C.I. = 0.0689, 0.1717), Diego Garcia and Salomon ($p = <0.0001$, 95% C.I. = 0.0570, 0.2082) and Peros Banhos and Egmont ($p = 0.0231$, 95% C.I. = 0.0066, 0.1244).

Tukey's HSD Test for multiple comparisons found that the mean value of body depth was significantly different between Diego Garcia and all locations. Diego Garcia and Egmont ($p = 0.0054$), Diego Garcia and Peros Banhos ($p = 0.0008$) and Diego Garcia and Salomon ($p = 0.0014$). No other significant differences were calculated. (Table A3.7, Figure 3.8).

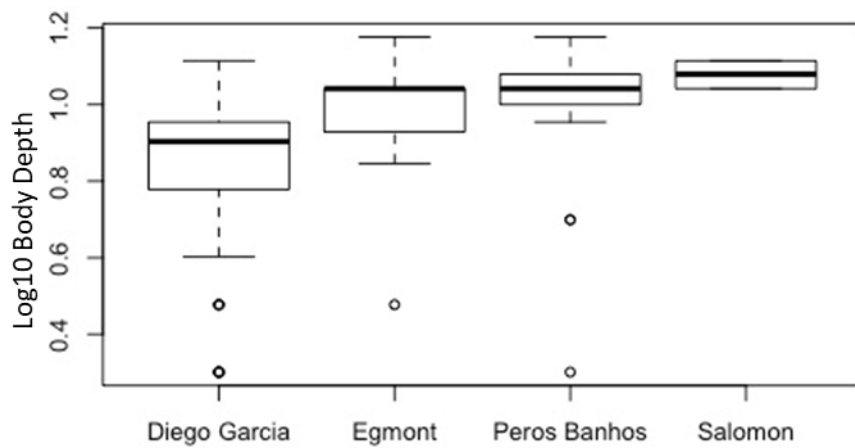


Figure 3.8 Box plot showing how average log body depth differs between sites for *L. gibbus*. ANOVA on circumference yielded significant variation between sites ($p = <0.0001$). Tukey HSD indicates significant differences for body depth between Diego Garcia and Egmont ($p = 0.0054$, 95% C.I. = 0.0384, 0.2957), Diego Garcia and Peros Banhos ($p = 0.0008$, 95% C.I. = 0.0630, 0.3078) and Diego Garcia and Salomon ($p = 0.0014$, 95% C.I. = 0.0831, 0.4467).

Tukey's HSD Test for multiple comparisons found that the mean value of fork length was significantly different between Diego Garcia and all locations. Diego and Egmont ($p = 0.0059$, 95% C.I. = 0.0152, 0.1213), Diego Garcia and Peros Banhos ($p = <0.0001$, 95% C.I. = 0.0611, 0.1619) and Diego Garcia and

Salomon ($p = <0.0001$, 95% C.I. 0.0649, 0.2153). No other significant differences were calculated. (Table A3.8, Figure 3.9).

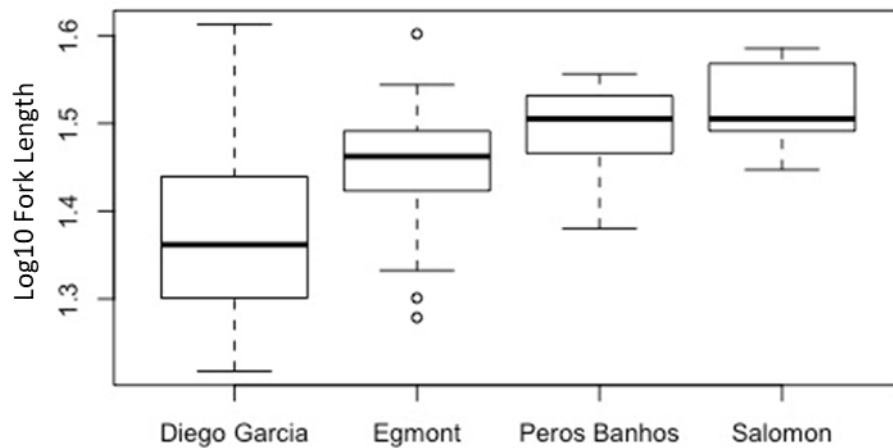


Figure 3.9 Box plot showing how average log fork length differs between sites for *L. gibbus*. ANOVA on fork length yielded significant variation between sites ($p = <0.0001$). Tukey HSD indicates significant differences for fork length between Diego and Egmont ($p = 0.0059$, 95% C.I. =0.0152, 0.1213), Diego Garcia and Peros Banhos ($p = <0.0001$, 95% C.I. 0.0611, 0.1619) and Diego Garcia and Salomon ($p = <0.0001$, 95% C.I. 0.0649, 0.2153).

3.3.6 Correlation test to measure the relationship between size variables

Pearson's correlation tests showed that the measured variables all have a significant positive linear correlation with one another for both species. The correlation matrix for *L. bohar* and *L. gibbus* size variables, and associated p values, are shown in Table 3.6 and Table 3.7 respectively. Size variables are all strongly correlated with one another for *L. bohar*, with high Pearson coefficients and p values less than 0.0001.

For *L. gibbus*, all size variables are strongly correlated, and the relationship is statistically significant, with p values of less than 0.0001, except for the correlation between body depth and weight. This correlation has a lower Pearson's coefficient compared to all other pairwise correlations, and a bigger p value which is less than 0.05, showing that this positive correlation is less statistically significant than all other comparisons. This suggests that for *L. gibbus*, body depth is not so greatly correlated with weight, compared with other size variables.

Table 3.6 Pearson's correlation analyses between log transformed variables for *L. bohar*. Values show Pearson's r value and associated significance level – $p < 0.05 = *$, $p < 0.001 = **$, $p < 0.0001 = ***$

	log Weight	log Circumference	log Body.Depth	Log Fork.Length
Logweight	1.0000	***	***	***
logCircumference	0.9088	1.000	***	***
logBody.Depth	0.8768	0.9542	1.000	***
logFork.Length	0.8992	0.9682	0.9689	1.000

Table 3.7 Pearson's correlation analyses between log transformed variables for *L. gibbus*. Values show Pearson's r value and associated significance level – $P < 0.05 = *$, $P < 0.001 = **$, $P < 0.0001 = ***$

	Log Weight	Log Circumference	Log Body.Depth	Log Fork.Length
Logweight	1.0000	***	*	***
logCircumference	0.8105	1.000	***	***
logBody.Depth	0.7130	0.6640	1.000	***
logFork.Length	0.7676	0.8884	0.7657	1.000

3.4 Results - shape variation

ANCOVA models demonstrate that location has a significant effect on the weight, circumference, and body depth of fish, even when controlling for fork length ($p < 0.0001$ for all size variables) and this initially indicates that there are differences in shape between sites (Table 3.8).

3.4.1 ANCOVA model to test for differences in shape of *L. bohar* between locations

The covariate, fork length, was significantly related to weight, $F(1,110)=426.53$, $p = <0.0001$). Location also had a significant effect on weight, after controlling for the effect of fork length, $F(3,110) = 20.97$, $p = <0.0001$). The ANCOVA results suggest that there are differences in weight between sites once fork length has been accounted for.

The covariate, fork length, was significantly related to circumference, $F(1,122)=1545.1$, $p = <0.0001$). Location also had a significant effect on circumference, after controlling for the effect of fork length, $F(3,122) = 124.6$, $p = <0.0001$). The ANCOVA results suggest that there are differences in circumference between sites once fork length has been accounted for.

The covariate, fork length, was significantly related to body depth, $F(1,123)=1541.6$, $p = <0.0001$. Location also had a significant effect on body depth, after controlling for the effect of fork length, $F(3,123) = 121.8$, $p = p = <0.0001$. The ANCOVA results suggest that there are differences in body depth between sites once fork length has been accounted for.

Table 3.8 Results for ANCOVA models controlling for the fork length in *L. bohar* at all sampling sites to determine whether there are differences in shape between locations. The model produced significant results for all comparisons when regressing log weight, log circumference, and log body depth against fork length between locations ($p = < 0.0001$).

Response Variable		Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	Location	3	1.097	0.366	20.97	8.1e-11 ***
	Fork Length	1	7.439	7.439	426.53	< 2e-16 ***
	Residuals	110	1.918	0.017		
Circumference	Location	3	0.3217	0.1072	124.6	<2e-16 ***
	Fork Length	1	1.3296	1.3296	1545.1	<2e-16 ***
	Residuals	122	0.1050	0.0009		
Body Depth	Location	3	0.3237	0.1079	121.8	<2e-16 ***
	Fork Length	1	1.3656	1.3656	1541.6	<2e-16 ***
	Residuals	123	0.1090	0.0009		

3.4.2 Residual analysis for *L. bohar*

Collinearity exists due to the high correlation between size variables (the pairwise correlations exceed the threshold of >0.7 for this dataset (Booth, Niccolucci and Schuster, 1994) – see Table 3.6). The ANCOVA model assumes that the response variable is independent of the covariable (fork length). The residual analysis model employs a sequential regression approach to calculate the independent contribution of each explanatory variable (Graham, 2003) (Table 3.9).

Residuals were extracted from the regression of log weight, log circumference, and log body depth on log fork length and put in an ANOVA against location to determine if the differences in weight, circumference and body depth between locations were statistically significant when collinearity was accounted for. Residual analysis suggested that there was no difference in weight when length is

adjusted for within the dataset ($F(3,111) = 1.566$, $p = 0.202$). Residual analysis also showed no difference in circumference when length is adjusted for ($F(3,123) = 1.878$, $p = 0.137$), or in body depth when length is adjusted for ($F(3,124) = 0.383$, $p = 0.765$).

Table 3.9 Residual analysis to evaluate the independent contribution of response variables to determine whether differences in weight, circumference and body depth were statistically significant for *L. bohar* when body depth was accounted for and collinearity effects removed.

Response Variable		Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	Location	3	0.0813	0.0271	1.566	0.202
	Residuals	111	1.9200	0.0173		
Circumference	Location	3	0.0048	0.0016	1.878	0.137
	Residuals	123	0.1052	0.0009		
Body depth	Location	3	0.0010	0.0003	0.383	0.765
	Residuals	124	0.1091	0.0009		

3.4.3 ANCOVA model to test for differences in shape of *L. gibbus* between locations

ANCOVA models demonstrate that location has a significant effect on the weight, circumference and body depth of fish, even when controlling for fork length ($p < 0.01$ and this initially indicates that there are differences in shape between sites (Table 3.10).

The covariate, fork length, was significantly related to weight, $F(1,85)=28.85$, $p < 0.001$). Location also had a significant effect on weight, after controlling for the effect of fork length, $F(3,85) = 4.018$, $p = 0.01$). The ANCOVA results indicate that there are differences in weight between sites once fork length has been accounted for.

The covariate, fork length, was significantly related to circumference, $F(1,95)=223.58$, $p < 0.001$). Location also had a significant effect on circumference, after controlling for the effect of fork length, $F(3,95) = 52.16$, $p < 0.001$). The ANCOVA results suggest that there are differences in circumference between sites once fork length has been accounted for.

The covariate, fork length, was significantly related to body depth, $F(1,95)=20.44$, $p < 0.001$). Location also had a significant effect on circumference, after controlling for the effect of fork length, $F(3,95) = 10.56$, $p < 0.001$). The ANCOVA results suggest that there are differences in circumference between sites once fork length has been accounted for.

Table 3.10 Results for ANCOVA models controlling for the fork length in *L. gibbus* at all sampling sites to determine whether there are differences in shape between locations. The model produced significant results for all comparisons when regressing log weight, log circumference, and log body depth against log fork length between locations ($p < 0.01$). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Response variable		Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	Location	3	0.2593	0.0864	4.018	0.01 *
	logFork.Length	1	0.6208	0.6208	28.851	6.7e-07 ***
	Residuals	85	1.8289	0.0215		
Circumference	Location	3	0.2921	0.0974	52.16	<2e-16 ***
	logFork.Length	1	0.4174	0.4174	223.58	<2e-16 ***
	Residuals	95	0.1774	0.0019		
Body depth	Location	3	0.9400	0.3133	10.56	4.66e-06 ***
	logFork.Length	1	0.6063	0.606	20.44	1.78e-05 ***
	Residuals	95	2.8181	0.0297		

3.4.4 Residual analysis for *L. gibbus*

As with *L. bohar*, collinearity exists due to the high correlation between size variables for *L. gibbus* (Table 3.7). Residuals were extracted from the regression of log weight, log circumference, and log body depth on log fork length and put in an ANOVA against location to determine if the differences in weight, circumference and body depth between locations were statistically significant when collinearity was accounted for. Residual analysis suggested that there was no difference in weight when length is adjusted for within the dataset ($F(3,86) = 0.111$, $p = 0.954$). Residual analysis also showed no difference in circumference when length is adjusted for ($F(3,96) = 1.559$, $p = 0.204$), or in body depth when length is adjusted for ($F(3,96) = 1.286$, $p = 0.284$).

Table 3.11 Residual analysis to independent contribution of response variables to determine whether differences in weight, circumference and body depth were statistically significant for *L. gibbus* when body depth was accounted for and collinearity effects removed.

Response variable		Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	Location	3	0.0071	0.0024	0.111	0.954
	Residuals	86	1.8290	0.0213		
Circumference	Location	3	0.0087	0.0029	1.559	0.204
	Residuals	96	0.1782	0.0019		
Body depth	Location	3	0.1146	0.0382	1.286	0.284
	Residuals	96	2.8532	0.0297		

3.5 Discussion

This research sought to explore whether any significant morphological variation occurred between sampling sites in the Chagos Archipelago. Whilst variation in average size between sites was shown to be significant, difference in shape was not significant when taking collinearity between the variables into account. The significant differences in size variables between atolls, for both species, are almost entirely between Diego Garcia and the other three atolls. Little significant difference in size was found between Egmont, Peros Banhos and Salomon. These findings suggest that variations in abiotic and biotic factors at Diego Garcia are driving the difference in lutjanid size.

There were some exceptions to the overall trend when looking at ‘weight’ as the size variable. For *L. bohar*, differences in weight between atolls was only significant between Diego Garcia and the two atolls of Egmont and Peros Banhos. For *L. gibbus*, weight was not statistically significant at all between any of the atolls. However, all other size variables (circumference, fork length and body depth) were significantly different between Diego Garcia and the other atolls, for both *L. bohar* and *L. gibbus*.

Fish species are highly plastic and exhibit a diversity of phenotypes in response to different environmental drivers and morphological differences arise as a result of this phenotypic plasticity (Olsson and Eklöv, 2005; Eklöv and Svanbäck, 2006; Chapman, 2015). Various environmental factors may be important for influencing phenotypic plasticity of fish and resulting morphological variations. These drivers of morphological variation include food availability (Abaad *et al.*, 2016), predation (Barley, Meekan and Meeuwig, 2017a; Hammerschlag *et al.*, 2018) habitat complexity, temperature, light (Kröger and Fernald, 1994), hydrodynamics (Collin and Fumagalli, 2011; Chapman, 2015), and prey biology (Davis, Pusey and Pearson, 2012).

Differences in the abiotic and biotic factors between Diego Garcia and the other atolls are explored more fully below.

3.5.1 Response to food availability

The impact on fish growth from reductions in ration size and feeding frequency on Lutjanus species has been well documented. Morphological differences in fish related to food availability and quantity have been described such that body size metrics (jaw depth, body depth, fork length) vary when food is scarce/abundant (Currens *et al.*, 1989; Abaad *et al.*, 2016; Jacobson *et al.*, 2018). Food availability for the two lutjanid species in this study (as measured by fish biomass) varies throughout the Chagos

Archipelago (Carlton, R., *et al.*, 2021) and differences in diet and nutrient availability between atolls may explain variations in size observed in our research.

Independent and extensive surveys of fish biomass were undertaken both prior to (2010 and 2012 (Graham *et al.*, 2013)), and at approximately the same time (2015) as the sampling undertaken for this research (Carlton, R., *et al.*, 2021). The surveys revealed that Diego Garcia had the lowest fish biomass, whilst the northern atolls of Salomon and Peros Banhos had much higher levels. The atoll of Egmont had the highest levels of fish biomass for the four atolls included in this research.

Surveys of differences in trophic levels and body size revealed that the highest proportion of larger fish (fork length greater than 30 cm) and those fish occupying higher trophic levels (trophic levels greater than 3.5) were found at the Egmont atoll. The fish assemblage at the northern atolls of Salomon and Peros Banhos, was similar to Egmont, although fish were marginally smaller. The surveys revealed that the smallest fish were found at Diego Garcia, and the proportion of those fish that occupied highest trophic levels were greatly reduced compared to the other atolls surveyed. These patterns of size distribution mirror the findings from our own research, where we also found the smallest individuals of both *L. bohar* and *L. gibbus* at Diego Garcia.

The results from the independent surveys, would also suggest that the relative abundances of the two focal species in our study would be reduced at Diego Garcia (compared to other atolls) because lutjanids are mesopredators, occupying high trophic levels greater than 4.0 (Froese R and D., 2017). However, relative abundances were not recorded for our research, so it is not possible to comment on this, other than to suggest reduced abundances of *L. bohar* and *L. gibbus* may be found at Diego Garcia compared to other atolls.

In addition to the abundance of food, the size of prey consumed by the fish can also affect morphology. Fish size has also been found to vary between locations for species where there are differences in prey size, such that fish have bigger gape, fork length and body depths where their prey items are bigger (Hegrenes, 2001; Paul *et al.*, 2017). Additionally, the nutritional value of the diet has also been shown to affect fish morphology. Lutjanids specifically, have been shown to grow bigger in terms of fork length and body depth, when they have access to a better diet, and in the absence of competition from other large piscivores / sharks (Barley., Meekan and Meeuwig, 2017b). These studies have shown that *L. bohar* and *L. gibbus* grow bigger and are in better condition when they consume a higher energy diet of fish and squid, compared to their conspecifics with a lower energy diet of benthic crustaceans and gastropods. A high energy diet also benefits an individual in terms of

its potential to pass on its genotype to subsequent generations because fish in good condition are more fecund, spawn earlier and produce bigger eggs which generate more viable larvae than those produced by fish in poorer condition (Ma, Kjesbu and Jørgensen, 1998; Marteinsdottir and Steinarsson, 1998). Genetic studies show that the fish communities of oceanic coral reefs are primarily self-seeding (Underwood, Travers and Gilmour, 2012). It is therefore likely that improved condition of *L. bohar* and *L. gibbus* at Egmont and the northern atolls may directly contribute to the greater abundances of offspring derived from these individuals in subsequent generations across the Chagos Archipelago.

The recreational fishery at Diego Garcia is most likely to be a significant factor contributing to the smaller size of lutjanids at this atoll. Whilst the fishery at Diego Garcia is subject to restrictions with regards to catch quotas (which permit pelagic and demersal fishing within less than 3 nautical miles of the atoll, such that any fish caught must be consumed within 3 days (Robert, 2007; MRAG, 2014)), anecdotal evidence indicates that these regulations are not adhered to. Fishing at Diego Garcia appears to have impacted fish biomass and fish assemblage at this atoll, thus reducing food availability for all species at this site (Meeuwig 2020, personal communication, Graham *et al.*, 2013).

Our findings indicate that the two focal species in our study at Diego Garcia are subject to factors that prevent them from growing as big as their conspecifics at other atolls (Egmont, Peros Banhos and Salomon) in the Chagos Archipelago. Differences in food resource availability and quality of diet are most likely a contributing factor to size differences between the atolls, with individuals at Egmont and the northern atolls having better access to food than fish at Diego Garcia. Gut contents analysis of individuals may help identify whether larger individuals at Egmont and the northern atolls are feeding on a higher energy diet than the lutjanids at Diego Garcia where fish stocks have been depleted.

3.5.2 Direct fishing pressure

In addition to recreational fishing at Diego Garcia reducing food availability (as measured by fish biomass), there is evidence that fishing could be directly impacting lutjanid size in Diego Garcia. Bigger fish are more vulnerable to exploitation and tend to be fished in preference to smaller individuals (Dulvy *et al.*, 2004). Fishing related depletion is particularly focussed on fish bigger than 50 cm and fish with trophic levels greater than 3.5 (Graham *et al.*, 2013). These characteristics fit the profile of mesopredatory species such as lutjanids. Snappers are typically targeted by fisheries, so it is highly likely that larger individuals of *L. bohar* and *L. gibbus* have been directly targeted in the recreational fishery area. The smaller size-class of individuals of both snapper species found at Diego Garcia is

most likely the result of over-exploitation and the lack of opportunity for individuals to reach full adult size, before being fished.

3.5.3 Predation risk

Shark abundance varies throughout the MPA and is greatest at Egmont, relatively high at the northern atolls, and lowest at Diego Garcia (Meeuwig, 2021, personal communication). The regions where shark biomass is greatest are also where the largest lutjanids are found (Egmont, Peros Banhos and Salomon). There are various factors that minimise the risk of predation for fish and these include fish behaviour (e.g. foraging on less risky food items and in areas where sharks/predators are less prevalent, and increased vigilance), fish size (bigger fish are less vulnerable to predation) and body shape (larger eyes, bigger caudal fins and narrower body shape that facilitates fast escapes from predators) (Dulvy *et al.*, 2004). Studies in the Indian Ocean have investigated the effect of shark predation on lutjanid morphology specifically and focussed on changes in body size/shape that are associated with predator depletion. These studies found that lutjanids were narrower and less deep-bodied for a given length (Barley, S., Meekan and Meeuwig, 2017) and had larger eyes/fins (Hammerschlag *et al.*, 2018) where shark biomass was high. In the first study the authors suggest that body shape could be attributable to poorer benthic diet in high predation environments and could also be an adaptation to acquire a body shape best suited to escaping predation. The second study also suggests that the morphology of large eyes and big caudal fins are morphological responses to high predation risk. However, whilst both these studies suggest that predators are driving the observed phenotypes, a genetic basis was not identified, and the authors were unable to determine whether these differences were adaptive or plastic genetic responses.

Fish size is important in the context of vulnerability to predation. Studies show that reef sharks are gape-limited and tend to feed on prey that are less than 36% of their length (Bethea, Buckel and Carlson, 2004). It is estimated that *L. bohar* enters size refuge at a length of approximately 55cm (Barley,S., Meekan and Meeuwig, 2017a; Barley,S., Meekan and Meeuwig, 2017b), and fork lengths of approximately 60 cm are common for long-lived adults (Marriott, Mapstone and Begg, 2007). Once they reach this length, they will enter size refuge from predation by sharks and will function as mesopredators (Barley, Meekan and Meeuwig, 2017b; Hammerschlag *et al.*, 2018).

With regards to *L. bohar* in this study, there was a relatively high proportion of individuals with fork lengths greater than the gape-limited size of 50-55 cm, at Egmont, Peros Banhos and Salomon (32 individuals from the whole dataset of 131). Whilst at Diego Garcia, there were only 2 fish exceeding

the shark gape-limit. These large individuals are most likely the product of good food availability, a high energy diet and a lack of competition for food resources at Egmont and the northern atolls.

The other lutjanid in this study, *L. gibbus*, is a smaller fish than *L. bohar* and average maximum fork length at adulthood is only 50 cm. This means that even if individuals survive into adulthood without being predated, and attain maximum size, most individuals will still be too small to enter gape-limited size refuge, and thus avoid predation by reef sharks. This therefore leaves *L. gibbus* more vulnerable to predation than *L. bohar*, particularly in Egmont, where shark biomass is greatest (Meeuwig, J, *et. al.* personal communication 2020). However, smaller fish can also be less profitable in terms of the energy derived from them when consumed, in comparison to larger prey, and small size can make fish less attractive to predators ((Catano *et al.*, 2016).

By contrast to the study by Barley *et al.*, we did not find significant differences in lutjanid shape for both species between atolls, so there is no evidence of predation risk driving differences in shape. We also found that *L. bohar* and *L. gibbus* were biggest at the atolls with the highest shark abundances, which suggests that the lutjanids in our study were able to access a high energy diet, even in the presence of other competitors / sharks. These findings suggest that the presence of sharks is not affecting lutjanid morphology at the reefs across the Archipelago, rather that access to a high energy diet is increasing body size at the atolls that are not affected by fishing pressure.

3.5.4 Habitat complexity

Habitat complexity can also contribute to phenotypic plasticity and differences in morphology between individuals of the same species occupying different habitats (Olsson and Eklöv, 2005) such that fish adopt the most optimal shape for foraging, swimming and predator avoidance for their particular environment. For example, fish that occupy pelagic waters may have longer, slimmer body shape to facilitate fast swimming, whereas conspecifics in more complex benthic / reef habitats have deeper bodies, larger heads and bigger gapes (Mihalitsis and Bellwood, 2019). Diego Garcia has the lowest structural complexity compared to the other atolls in this study, with the northern atolls of Peros Banhos and Salomon having highest structural complexity (Graham *et al.*, 2013; Carlton, R.*et al.*, 2021). Whilst differences in reef structural complexity may contribute towards the different size morphologies of lutjanids we observed in our study, it is difficult to isolate the strength of this as an influence, because our size metrics only recorded variables such as fork length and body depth – but not gape, head size, or caudal fin height (which are additional measurements used to detect phenotypic change in response to habitat complexity). Additionally, even if variables such as body

depth and fork length are used as a proxy for investigating if fish are showing differences in morphology associated with habitat complexity, caution must be taken in interpreting our results. Although we detected smaller body depths and shorter fork lengths for lutjanids at Diego Garcia compared to the other three atolls in our study, these values decrease proportionately with one another, and shape is not changing for the fish in this study. For example, fish are bigger for all variables at the northern atolls and Egmont, compared to Diego Garcia, and if body depth is greater, so is fork length – consequently, fish do not appear to have a morphology that reflects acclimatisation to a more complex habitat at Peros Banhos, Salomon and Egmont.

The structural complexity of reefs at each atoll may also influence the risk of predation to smaller individuals as the reef habitat, and particularly the hard coral architecture, can provide refuge for prey (Madin, Madin and Booth, 2011). The reefs at Egmont, Peros Banhos and Salomon are more diverse in terms of hard coral cover and are more structurally complex than the atoll at Diego Garcia (Graham *et al.*, 2013; Carlton, R., *et al.*, 2021). Although the risk of predation for *L. gibbus* is reduced at Diego Garcia, the atolls of Egmont, Salomon and Peros Banhos might be more attractive for this smaller lutjanid, due to the higher fish biomass (and food availability) and the protective refuge provided by the more complex coral reef habitat.

3.5.5 Temperature

Another factor that can influence phenotypic plasticity of fish is water temperature, which has been shown to affect fish during early life, and influences growth and survival (Crozier and Hutchings, 2014; Jonsson and Jonsson, 2014; Vagner, Zambonino-Infante and Mazurais, 2019). Fish are poikilothermic and are unable to regulate their internal body temperature – as such, it alters according to environmental temperature. Therefore, environmental temperature is important and operates within optimal ranges for fish, affecting metabolic rates, feeding activity, maturation, distribution and survival (Hochachka, Peter and Somero, 2002). Divergence from optimal temperature ranges during the early stages of fish development can affect these physiological and developmental processes, leading to extensive variation in phenotype (Johnston and Bennett, 2008).

Fish traits affected by temperature include body shape (Ramler *et al.*, 2014), size (Jonsson and Jonsson, 2014), meristic traits (e.g. number of fins (Christou *et al.*, 2018)), growth rate (Crozier and Hutchings, 2014), lifespan (Lee, Monaghan and Metcalfe, 2013), muscle structure (Albokhadaim *et al.*, 2007) and stress response (Mateus *et al.*, 2017). Although our study did not detect differences in body shape between lutjanids in the Chagos Archipelago, the size differences that we detected could be

worth exploring further in the context of sea temperature differences. Size could indicate differences in early stages of growth as rapid early growth has been associated with increased water temperatures (Jonsson and Jonsson, 2014).

Globally, sea surface temperatures increased sharply during 2015 and 2016 as a consequence of the El Nino Southern Oscillation climatic event, which lead to mass coral bleaching (Head, *et al.*, 2019). Within the Chagos Archipelago there were differences in the geographical distribution and levels of thermal stress caused, with the warming event of 2016 being more extreme and thermal anomalies were greatest in the south-west of the Archipelago (the region encompassing Egmont and part of the Great Chagos Bank). The back to back sea-warming events of 2015 and 2016 caused a 60% decline in hard coral cover, with most coral loss occurring in 2015 (Head, *et al.*, 2019).

The effects of the two climatic events at the Chagos Archipelago could be two-fold: firstly, the loss of such a high proportion of hard coral structure across the site could mean a decrease in structural complexity at specific atolls, which has also been proposed as a potential driver of phenotypic change in fish (Olsson and Eklöv, 2005; Mihalitsis and Bellwood, 2019). There was extensive hard coral mortality throughout the Chagos Archipelago and the surveys found that community composition of hard corals within the region shifted (Head *et al.*, 2019). However, reef structural complexity, as opposed to coral cover, is the important driver for the processes driving altered fish morphology (Olsson and Eklöv, 2005), and survey data for structural complexity is not available for the 2015/2016 coral bleaching events. Whilst it is possible that coral mortality may have caused loss of structural complexity at localised sites, without empirical data it is impossible to draw any conclusions around this and any impact it may have had on fish phenotype.

The second factor affecting fish populations in the Chagos Archipelago following the global climatic events of 2015/16, is the increase in sea surface water temperature. Whilst the thermal anomalies were more extreme in the region surrounding Egmont in 2016, the effects of increasing sea surface temperature were experienced throughout the whole Archipelago, with temperatures increasing at all locations. It is likely that the increase in sea surface temperature could drive phenotypic changes in fish across the Archipelago.

The samples for this study were collected in November 2015 and May 2016 – so sample collection coincides with the timings of the two climatic events. Although the precise ages of the fish in our study are unknown, their sizes are representative of juveniles (2-3 years old) through to fully grown

adults (Fortaleza and Nanola, 2017; Moore, 2019). Studies have shown that other lutjanid species can reach fork lengths of approximately 25cm (which equates to the smaller individuals in our study) at approximately 18 months, so this is the minimum age that individuals in our study are likely to be. The effect of temperature increases on fish size, growth and rate of growth primarily causes effects in juveniles and adults and is due to temperature increases at the embryonic / or larval phase. Warmer sea temperature can increase the rate of growth in juveniles such that they reach maturity quicker and at smaller sizes (Kausar and Salim, 2006; Jonsson and Jonsson, 2014). However, it is not likely that the fish in our study were at the early developmental stage where the effects of temperature on phenotype would be acting, as they would be at the older end of the juvenile age range, or adults. It is possible that some of the younger fish that hatched during 2015 were in the earlier developmental stages during the first climatic event in 2015, but as we have only individual time point metrics for size variables, it is not possible to calculate rates of growth (which may vary in response to sea surface temperature warming). It is likely that the phenotypes of the fish in our study were the product of the environment prior to the 2015 / 2016 climatic events.

Other factors that could affect phenotypic plasticity between atolls, such as, light levels, and current velocity were not measured. Light levels (in combination with habitat structure and complexity) influence pigmentation in fish and are associated with predation risk for individuals as light levels vary between locations (Gusen, 2010; Mitchem *et al.*, 2018). As the atolls we studied occupy similar oceanographic settings, it is likely that they are subject to similar light regimes. However, as we did not investigate pigmentation of the fish, light level is not a factor that can be considered in the context of our study.

3.6 Conclusion

Our study has identified differences in lutjanid body size such that individuals are smaller at the fish-depleted atoll of Diego Garcia, compared to other atolls in the Chagos Archipelago. In a natural study environment, it is difficult to isolate the specific environmental driver, or combination of factors, that give rise to the differences in size of *L. bohar* and *L. gibbus*, between Diego Garcia and the other atolls we studied. It seems likely that the primary driver of the morphological size differences we observed, is accounted for by variations in food availability and nutritional value of food resources, between the atolls. Several factors appear to be driving the diet related variation between the atolls. Fishing at Diego Garcia has diminished fish biomass, and therefore available food. A second factor affecting fish biomass at Diego Garcia is the lower levels of reef structural complexity at this location, compared to the other atolls we sampled fish from. The increased fish biomass at Egmont and the northern atolls

suggests that lutjanids at these sites have access to a better, higher energy diet compared to individuals at Diego Garcia and therefore, grow bigger. Fishing at Diego Garcia may also be affecting the lutjanid assemblage at Diego Garcia directly, due to the fishing of larger individuals, thus removing them from the population.

Analysis of gut contents for individuals at each location could provide useful information regarding differences in diet between Diego Garcia and the other atolls and identify whether the bigger fish are eating a more energetically rewarding diet. If *L. bohar* and *L. gibbus* at Egmont, Salomon and Peros Banhos benefit from a higher energy diet, they will also be in better condition and therefore more likely to pass on their genotypes to subsequent generations. Conversely, across the Archipelago as a whole, there could be a decrease in abundance of those individuals (and their associated genotypes) derived from the lutjanids at Diego Garcia who are in worse condition due to the lower energy diet at this atoll.

This study cannot distinguish the precise mechanisms underlying the morphological variation between populations of fish at the atolls of the Chagos Archipelago and the extent to which phenotypic plasticity in response to different environments plays a role. Additional experimental and genetic data could supply some of this information – for example, analysis of population structure between locations may detect structuring related to the accumulation of those individuals with genotypes that are favourable for the environmental conditions in the system. Analysis of the gene expression profile of individuals at each atoll might also provide information regarding the genetic mechanisms driving phenotype in response to environmental differences – for example RNA sequencing can provide a functional profile of gene expression in specific tissues, to identify upregulation / down-regulation of genes controlling traits of interest. Studying the two focal species over a longitudinal time frame would also be beneficial as there is no historical data available regarding lutjanid morphology, so it is impossible to draw conclusions regarding how this may have changed over time.

The null hypothesis that there is no significant difference in shape between sites within the Chagos Archipelago can be accepted. However, the null hypothesis that there is no difference in size for *L. bohar* and *L. gibbus* between sites, can be rejected. These findings indicate that even in the relatively untouched location of the Chagos Archipelago, reef fish are still subject to the anthropogenic impacts of fishing that can cause variations in morphology.

Chapter 4 - Population Genetic Structure of *L. bohar* and *L. gibbus* in the Chagos Archipelago using RAD-sequencing

4.1 Introduction

Population genetics has been applied to multiple studies investigating marine species that are subject to exploitation to understand the various drivers that impact population structure and enable good conservation and fisheries management. Analysis of population genetic structure of species facilitates the calculation of gene flow and connectivity between populations, as well as identifying barriers to exchange of individuals. Population connectivity is important for all organisms, even when populations are independently sustainable, as good connectivity promotes gene flow and associated genetic buffers that facilitate a response to environmental challenges, and recovery following local extinctions (Allison *et al.*, 2003). Connectivity between marine populations at different locations is important for protection and conservation of marine biodiversity, particularly for exploited marine taxa. Understanding the factors contributing to population structure for marine species can also be used to inform the definition and range of protected areas (Allison *et al.*, 2003). Genetic differences accumulate when populations become reproductively isolated. Therefore, a spatial analysis of genetic differentiation provides an overview of population connectivity and underlying processes driving these patterns (Hedgecock *et al.*, 2007).

Marine fish populations have been considered panmitic and well connected, due to long pelagic larval dispersal potential and a perceived lack of physical barriers to dispersal in the oceans, which can lead to a lack of population structure over a geographical scale (Palumbi, 1994). However, this perception is not applicable in all cases, with population structure being exhibited over a range of geographic scales for some species (Knutsen *et al.*, 2011; Underwood, Travers and Gilmour, 2012). Multiple factors may contribute to the genetic subdivision evident in these examples, including lower pelagic larval durations, larval behaviour, 'sweepstakes reproductive success', and local oceanographic conditions and processes (Underwood, Travers and Gilmour, 2012; Puritz, Gold and Portnoy, 2016).

Lutjanid species have a global distribution and are widely exploited throughout their range. Whilst various population level studies have been conducted on other marine fish, including lutjanid species, most have been carried out in areas subject to multiple anthropogenic impacts and where overfishing is common or has been occurring at high levels historically (Zhang, Cai and Huang, 2006; da Silva *et al.*, 2016). The atolls of the Chagos Archipelago provide a natural model system to investigate population connectivity. Due to their remote location and MPA status they are relatively unaffected

by anthropogenic impacts, except for illegal shark fishing, recreational fishing and bleaching of corals due to climate warming (Sheppard *et al.*, 2012; Graham *et al.*, 2013). The atolls are geographically isolated and separated by long stretches of deep ocean from other reef locations (Cocos Keeling Islands are 2750 km East and nearest shallow water reefs west are Saya de Malha, 1050 km away). The remote location and reduced anthropogenic influences present at the Chagos atolls allow for the study of population connectivity of lutjanids in a relatively unspoilt environment.

The two focal lutjanid species investigated in this study – *Lutjanus bohar* and *Lutjanus gibbus* – are distributed throughout the Indo-Pacific Oceans and are commonly fished in much of their range where they encounter human populations. They are large mesopredatory teleosts inhabiting coral reefs, sheltered lagoons and outer reefs, and exhibit biological attributes in common with other lutjanid species, including a relatively long pelagic larval phase that has been estimated to be upwards of the mean duration for lutjanidae of 30 days (Allen, 1985). The long pelagic dispersal duration for lutjanids provides the potential for dispersal over long distances and is combined with adult migrations of the order of 100s of km. In the absence of other barriers to dispersal, genetic connectivity between populations would be anticipated to be high over small geographical scales, such as the region covered by the Chagos Archipelago, where atolls are separated by distances of 50 – 200 km.

Other studies have looked at the population structure of Lutjanid species over relatively large geographic scales (i.e., within oceans) and findings indicate that populations are well connected and genetically homogeneous (da Silva *et al.*, 2016). Where studies have detected structure, this has been low and is evident where there are physical barriers – for example the Indo-Pacific barrier (Fleminger, 1986).

Other studies looking at population structure for lutjanids have used low numbers of molecular markers (including mitochondrial sequences and limited nuclear markers) to generate population level markers for genetic connectivity analysis (Muths *et al.*, 2012; da Silva *et al.*, 2016). This study used next generation RAD sequencing (RAD-seq) technology to generate large numbers of Single Nucleotide Polymorphisms (SNPs) as molecular markers for individuals of both species at each sampling site. SNP data was used to generate population structure level statistics to determine population differentiation and genetic structure between atolls.

The aim of this study is to explore the population structure and genetic connectivity of *L. bohar* and *L. gibbus* between the four atolls of Diego Garcia, Egmont, Salomon and Peros Banhos in the Chagos

MPA, using RAD-seq to generate multiple markers. Based on existing knowledge of population structure of other marine fish, the null hypothesis is that there is weak population structure between sites in the Chagos Archipelago and that populations are well connected.

4.2 Methods

4.2.1 Sample collection

Samples of *L. bohar* and *L. gibbus*, were collected from the Chagos Archipelago MPA during expeditions carried out in 2015 and 2016. In line with the requirements due to the protected status of the region, Foreign Commonwealth Office approval was secured for the research. Samples of both species were collected from four atolls distributed across the site: Diego Garcia, Egmont, Salomon Islands and Peros Banhos (Figure 4.1 and Figure 4.2). The atolls were separated by distances of approximately 50 – 200 km.

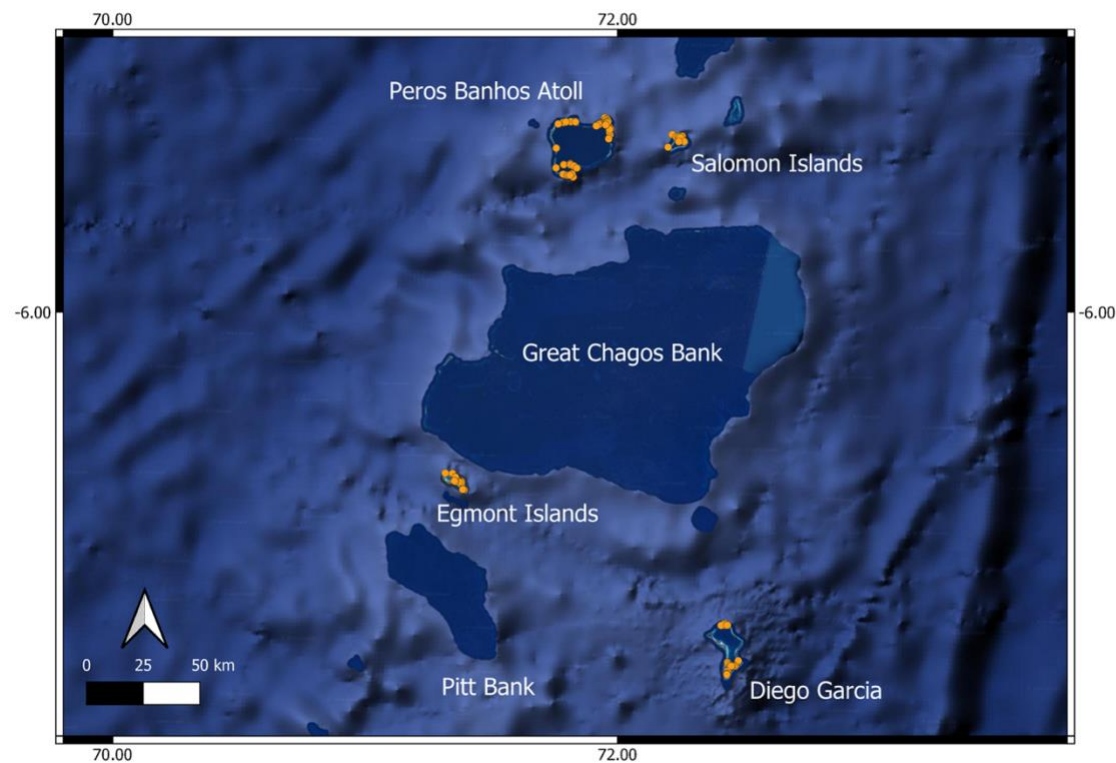


Figure 4.1. Sampling locations for *L. bohar* at atolls within the Chagos MPA (orange dots denote each sampling site)

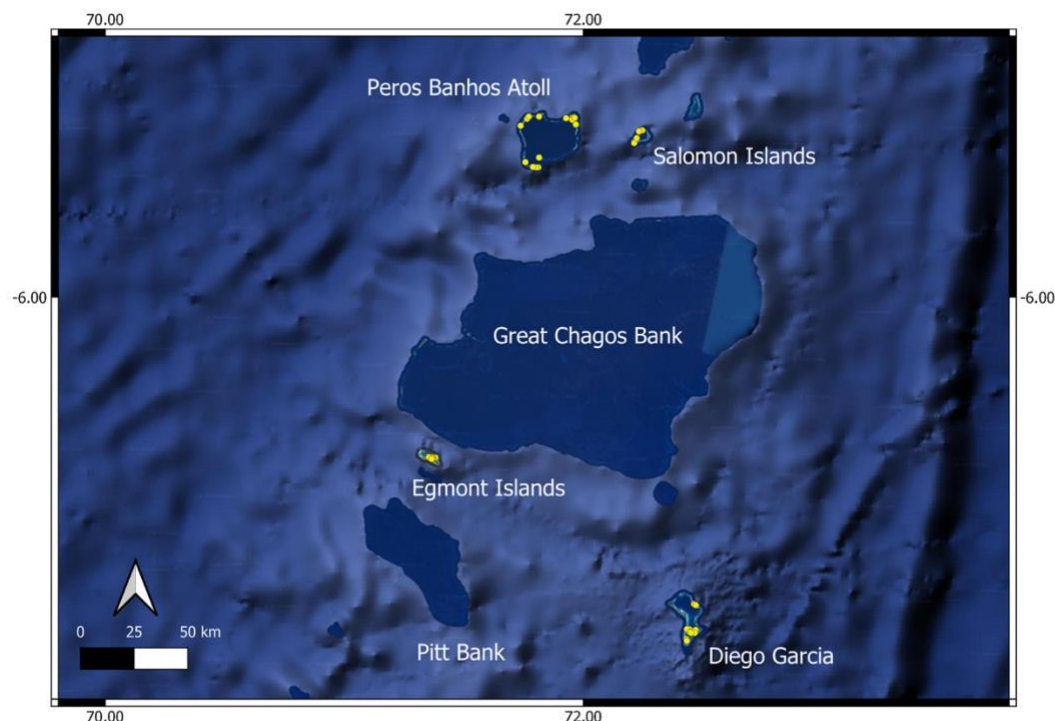


Figure 4.2 Sampling locations for *L. gibbus* at atolls within the Chagos MPA (yellow dot denotes each sampling location).

Adults of both species were captured using pole and line methods, by professional fishermen and subject to the necessary ethical approval and independent assessment by a licensed veterinarian. For individuals longer than 25 cm, a 0.5 cm³ punch biopsy of muscle tissue was taken from the ventral dorsal area of the fish and immediately stored in RNAlater™ (ThermoFisher Scientific) at 4°C until being transferred to -80 °C storage at the UK laboratory. Care was taken when collecting the tissue biopsy not to damage the fish or cause bleeding. Fish were subsequently tagged and released live.

4.2.2 DNA extraction

High molecular weight genomic DNA was extracted from approximately 20 mg of muscle tissue for each sample using the Qiagen Dneasy Blood and Tissue Kit (Qiagen, UK). Genomic samples stored in RNAlater™ were thawed on ice, then dissected using sterile, disposable scalpels to provide two sections and the dissected tissue sample was used for genomic DNA extraction.

The manufacturer's protocol was followed, with the addition of Riboshredder™ (Epicenter,US) following tissue lysis. Riboshredder™ is an RNase blend that degrades unwanted RNA in DNA purification protocols. Samples were eluted into 35 µL Qiagen AE elution buffer (10mM Tris-Cl, 0.5

mM EDTA, pH 9.0). Genomic DNA was successfully extracted from 167 *L. bohar* and 88 *L. gibbus* tissue samples.

To determine DNA concentration, a Qubit® fluorometric quantitation assay was used. This system employs technology to detect target-specific fluorescence and is more sensitive than conventional UV-absorbance based quantitation assays. The concentration of extracted DNA was determined using the Qubit® High Sensitivity assay, using a Qubit® 3.0 Fluorometer (Life Technologies).

To confirm sample quality, samples were run alongside a 1kb DNA Ladder (NEB Quickload®), on 1.2% agarose (Sigma) 0.5x TAE gel, using SYBR™ Safe (Invitrogen) DNA stain at 1:10,000 dilution. Samples with concentration less than 20 ng/μL and / or showed degradation in gel electrophoresis were discarded. The remaining samples were selected for RAD-sequencing, whilst also ensuring a representative distribution of samples based on geographical location (Diego Garcia, Egmont, Peros Banhos and Salomon).

Samples were normalised to a concentration of 20-30 ng/μL, as required by the sequencing facility, and 40 μL of each sample was loaded into a 96-well plate and sent to Floragenex (Oregon, US) on dry ice. Samples that did not meet this requirement, or that showed any degradation in gel electrophoresis, were discarded at this step to ensure that only samples of high quality were used for RAD sequencing. Following the quality control assessment, a total of 120 *L. bohar* samples and 70 *L. gibbus* samples were submitted in two separate batches (Run_1 and Run_2) to the sequencing facility for RAD-sequencing.

4.2.3 RAD-seq library preparation

RAD-libraries were prepared using methods described in (Etter *et al.*, 2011). Genomic DNA was size selected at 500-750 ng and digested with the restriction enzyme *Sbf1* (New England Biolabs). *Sbf1* recognises the restriction cut site 5'...CCTGC*AGG...3' and genomic DNA digestion generated restriction fragments 5-25 kb in size. Subsequently, modified Illumina P1 adapters containing a 10 bp unique barcode (for downstream sample identification) were ligated to restriction fragments. The resulting construct was randomly sheared with a Bioruptor (Diagenode) and resulting fragments of DNA size-selected (<300-500 bp). Following Illumina P2 adaptor ligation, each RAD fragment library was PCR amplified and purified.

The amplified PCR fragments were visualised on an agarose gel and an Agilent Bioanalyzer trace (Agilent, Santa Clara, US) used to validate library integrity. Quantitative PCR was run on samples to estimate optimal next-generation sequencing plating concentration.

The RAD fragment libraries comprising 95 samples and 1 internal control per 96-well plate, were allocated to a single lane of 2x100bp Illumina HiSeq 4000 machine (University of Oregon Genomics Core Facility), generating 90-130 million indexed reads per sample, with 80% of bases Q30 or better. Samples were sequenced from both ends of each sequence to generate paired end reads. Sequence data was delivered in FASTQ format, ready for bioinformatic analysis. It was anticipated that adopting this workflow RAD-sequencing would generate a minimum of 30,000 loci in fish (sufficient loci for RAD-seq population analysis) (Jackson *et al.*, 2014; Puritz, Hollenbeck and Gold, 2014).

4.2.4 RAD-tag data analysis

Generated RAD-tags were analysed using the pipeline developed in Stacks version 2/2.4.1 (Catchen *et al.*, 2013), comprising the components detailed in Appendix Table A4.1.

4.2.5 Demultiplexing of data – *process_radtags*

Quality filtering and demultiplexing was carried out using the *process_radtags* module under default parameters to clean data by removing reads with ambiguous barcodes, as well as removing any reads with an uncalled base, and discarding reads with low quality scores. The data was filtered to ensure that only individuals with more than 900,000 retained reads were kept (2 *L. bohar* and 1 *L. gibbus* samples were discarded). Samples were then separated into separate directories according to species (*L. bohar* and *L. gibbus*) for subsequent downstream assembly of loci.

4.2.6 Parameter Selection – *denovo_map.pl*

Several parameters determine how loci are assembled and aligned, and which reads form a locus. Parameter choice controls stack assembly and is determined by biological features of the RAD-seq data (such as inherent polymorphism within the samples), restriction enzyme used (which controls number of cut sites), length and number of raw reads and degree of coverage. The number and polymorphism of loci in a RAD-seq dataset, generated in the Stacks *denovo* assembly pipeline (*denovo_map.pl*) are determined by three main parameters which control the formation of putative alleles (stacks) and the degree of mismatch permitted when merging stacks and creating the catalog of loci: *m*, *M* and *n*; This is detailed in Appendix Table A4.2.

In order to make an informed choice regarding optimal combination of parameters for the dataset, the protocols adopted by Rochette and Catchen (2017) and Paris *et al.* (2017) were followed.

Briefly, the raw results of the pipeline were then filtered using the *populations* component, retaining only loci shared by at least 80% of samples for both species (*r 0.80 loci*) (see Appendix Table A4.3, Table A4.4). To select the best parameter set to apply to the data for downstream analysis, the number of polymorphic loci shared by 80% of samples and the distribution of number of SNPs per locus were analysed (see Appendix Figure A4.1, Figure A4.3)

The optimal value for the number of raw reads used to create a putative allele or stack (*m*) was set at *m* = 3, as this was previously demonstrated to provide polymorphism across a range of datasets from different species, including fish (Paris *et al.*, 2017). The optimal value for mismatches allowed between stacks (*M*) was inferred by taking the value at which the number of shared polymorphic loci plateaus and the value at which loci with 1-5 SNPs begins to stabilise. For both the *L. bohar* and *L. gibbus* dataset, this value was between 4-5 (see Appendix Figure A4.2, Figure A4.4).

The optimal value for the mismatches allowed between loci during formation of the stack catalog (*n*) was subsequently inferred using the previously determined values for *m* and *M*, following the protocol in Paris *et al* 2017 and the *denovomap.pl* component of Stacks. Numbers of assembled loci, polymorphic loci and SNPs were calculated for the *r80 loci* for each output data set. A value of *n* = 4 gave the highest number of polymorphic loci for both datasets (see Appendix Table A4.5).

For both datasets, following parameter optimisation, RAD loci were formed allowing a maximum of five nucleotide mismatches (*M* = 5), minimum depth of coverage to create a stack of *m* = 3, and number of mismatches allowed between sample putative loci when generating the catalog of *n* = 4 (Table A4.2). Reads were aligned *de novo* with each other to create a putative catalogue of RAD loci, employing the *cstacks* module in Stacks.

4.2.7 Population analysis using *Stacks populations* bioinformatics pipeline

For both datasets, a population map was supplied to the *populations* component of Stacks, assigning samples to their original sampling location: Diego Garcia, Salomon, Egmont or Peros Banhos.

An array script was executed to calculate iterations of -p (the minimum number of locations a locus must be present in order to be processed) and -r (the minimum percentage of individuals a locus must be present in to be processed). Values for -p ranged from 1-4 and for -r from 10%-95%. This generated retained SNPs genotyped in at least 10 to 95% of individuals at 1 to 4 sampling locations.

Population level F statistics including π (nucleotide diversity), F_{IS} (inbreeding coefficient), H_E (expected heterozygosity) and H_O (observed heterozygosity) were computed. To detect whether genetic structure between sampling locations existed, pairwise F_{ST} comparisons were computed using the *populations* module of Stacks. First, markers that were out of Hardy-Weinberg Equilibrium (HWE) (P-value 0.05) at all locations were removed. Polymorphic loci with a minor allele frequency (MAF) > 0.05 across all sampling locations were kept, in order to eliminate biases in genetic connectivity associated with retaining low frequency SNPs (Roesti, Salzburger and Berner, 2012). F_{ST} values were calculated with a p-value correction of 0.05 and a genome-wide Bonferroni correction. Files containing genotype data in *Structure* and *Genepop* formats were generated for subsequent genetic diversity and population structure analysis.

4.2.8 PCA Analysis of population structure using adegenet for R

The adgenet package was implemented in R, to carry out a Principal Component Analysis (PCA) of the genotypes generated from the Stacks *populations* output. PCA enables the visualisation of multivariate data by transforming multiple correlated quantitative variables into fewer dimensions or “principal components”(Lever, Krzywinski and Altman, 2017). Principal components maintain trends in data, whilst simplifying it. PCA is a commonly used treatment of high-dimensional data, (characteristic of genetic data), where there are multiple linked data for each sample analysed. Prior to PCA, the information in a dataset comprises the total variation contained within the dataset. PCA identifies the principal components contributing the most to the variation in the dataset and reduces the dimensionality of the data from multiple variables to two (or three) principal components which can be easily visualised as a graph.

PCA was carried out on the RAD-seq derived SNP dataset generated, to see whether there were patterns associated with genotype variation and geographic location for *L. bohar* and *L. gibbus*. PCAs were performed using the R package *adegenet* (Jombart and Ahmed, 2011), without any *a priori* grouping assumption and applying false discovery rate correction of 0.05. Using the *genepop* output datafile generated in Stacks, PCA was carried out on neutral SNPs (those alleles significantly at HWE) and outlier SNPS (those alleles not significantly at HWE). PCAs were carried out for each of the

iterations generated by the *populations* component of *Stacks* to look for differences in differentiation and clustering for each of these models.

4.3 Results

4.3.1 RAD-seq analysis – Running *process_radtags* to demultiplex raw data

Following cleaning to remove ambiguous RAD-tags and low-quality reads, 74.30% (717.8×10^6) of cleaned reads remained from plate 1 (Run_1), and 90.0% (627.0×10^6) of cleaned reads were retained from plate 2 (Run_2; Table 4.1).

Table 4.1 Percentage retained reads for run 1 and 2 following demultiplexing of raw RAD- sequences using *process_radtags*

	Run_1	Run_2
Total sequences	965976480	696310722
Ambiguous barcodes	197386826	60090500
Low quality	39232998	4810892
Ambiguous RAD-Tag	11581764	4440597
Retained reads	717774892	626968733
% Retained reads	74.31	90.04

All samples from Run_1 retained more than 900,000 reads and were selected for downstream analysis. The lowest number of retained reads for Run_1 was 2.09×10^6 and the maximum was 18.49×10^6 . The mean number of reads per sample was 7.35×10^6 . For Run_2, three samples retaining fewer than 900,000 reads were discarded from further analysis. The lowest number of retained reads for Run_2 was 1.59×10^6 , and the maximum was 35.89×10^6 .

In total, from 120 *L. bohar* and 70 *L. gibbus* sequences submitted for RAD-sequencing, 118 *L. bohar* and 69 *L. gibbus* sequences were retained for downstream analysis.

4.3.2 RAD-seq analysis – Genotyping datasets for population structure analysis

Following filtering applied in *denovomap.pl* ($M = 5$, $n = 4$, $m = 3$), the number of RAD loci per individual ranged from 20,939 to 83,267 for the *L. bohar* dataset (comprising 118 individuals), with average number of loci per sample being 61,312. The number of reads per sample post-filtering ranged from 363,745 to 8,238,457, with the average reads per sample being 2,800,560. Mean coverage per sample was between 17 to 100 x, with average coverage per sample being 44 x. For the *L. bohar* data set,

there are approximately 337,000 assembled loci, of which 112,216 are polymorphic loci with 2.23 million SNPs (see Appendix Table A4.6)

For *L. gibbus* (comprising 69 individuals), the number of RAD loci obtained per individual post filtering with *denovomap.pl*, ranged from 31,583 to 73,449, with an average of 61,886. The number of reads per sample ranged from 644,641 to 14 million, with an average of 3.6 million reads. Mean coverage per sample ranged from 20 to 199 x, with an average coverage per sample of 53 x. For the *L. gibbus* data set, there are approximately 313,000 assembled loci, of which 112,738 are polymorphic loci, comprising 1.79 million SNPs (see Appendix Table A4.6).

The number of loci and SNPs present in at least 75 to 95 % of individuals at all four sampling sites are shown in Table 4.2. However, subsequent population analysis focussed on SNPs present in most individuals (more than 75%) at all four sampling locations and for a minor allele frequency greater than 0.05. For *L. bohar*, this resulted in more than 20,000 SNPs being genotyped and for *L. gibbus*, 11,496 SNPs were genotyped (Table 4.2).

Table 4.2 Number of loci and SNPs present in at least 75 to 95% of *L. bohar* and *L. gibbus* at all four sampling locations, post filtering in populations, using the parameters $M=5/n=4/m=3$ in *Stacks*

% of individuals with locus	M=5/n=4: <i>L. bohar</i>		M=5/n=4: <i>L. gibbus</i>	
	Total SNPs	Total Loci	Total SNPs	Total Loci
75	20065	43999	11476	26816
80	13016	32150	6396	16746
85	7056	20441	1177	4110
90	1626	6669	428	1823
95	55	416	2	21

4.3.3 RAD-seq analysis - F_{ST} analysis of population structure

Population structure analysis included a total of 118 *L. bohar* and 69 *L. gibbus* individuals, distributed at the four different atolls in the study (Table 4.3). Pairwise F_{ST} comparisons revealed low levels of genetic differentiation across the study region for both species (Table 4.4).

Table 4.3 Numbers of *L. bohar* and *L. gibbus* individuals (n) at each sampling site used for population structure analysis and calculating F_{ST} (genetic differentiation) values

Location	<i>L. bohar</i> (n)	<i>L. gibbus</i> (n)
Diego Garcia	42	30
Egmont	17	15
Peros Banhos	45	12
Salomon	14	12
Total (n) across study region	118	69

Genetic structure was calculated based on loci being present in the majority of individuals (75%) at all four sampling locations. For both species, the greatest population structure is evident between the Northern atolls (Peros Banhos and / or Salomon) and the central atoll of Egmont. The lowest differentiation and therefore the greatest genetic connectivity exists between Diego Garcia and all other atolls.

4.3.4 Population Structure of *L. bohar* and *L. gibbus* using pairwise F_{ST} comparison

RAD-seq analysis of loci present in most individuals (more than 75%) revealed evidence of extremely weak population structure for *L. bohar* between all atolls (Table 4.4, Figure 4.3). The greatest differentiation exists between Salomon and Egmont ($F_{ST} = 0.0071$). The highest gene flow is between Diego Garcia and Peros Banhos ($F_{ST} = 0.0032$) and pairwise F_{ST} values between all other atolls are lower and similar to one another. (These findings of very weak differentiation between sites are also evident when looking at SNPs present in 80-95% of individuals – for further details refer to Appendix Table A4.7)

Table 4.4. F_{ST} values for pairwise comparisons between all four sampling sites for SNPs present in >75% of individuals for *L. bohar*. Parameter values for RAD-locus assembly were $M = 5$, $n = 4$ and $m = 3$. Highest level of statistically significant F_{ST} values is in bold. Bonferroni corrected p-value of < 0.05 applied to all values. n is number of individuals.

Species	n	SNPs		Egmont	Salomon	Peros Banhos
<i>L. bohar</i>	118	20,065	Diego Garcia	0.0039	0.0042	0.0032
			Egmont		0.0071	0.0045
			Salomon			0.0044

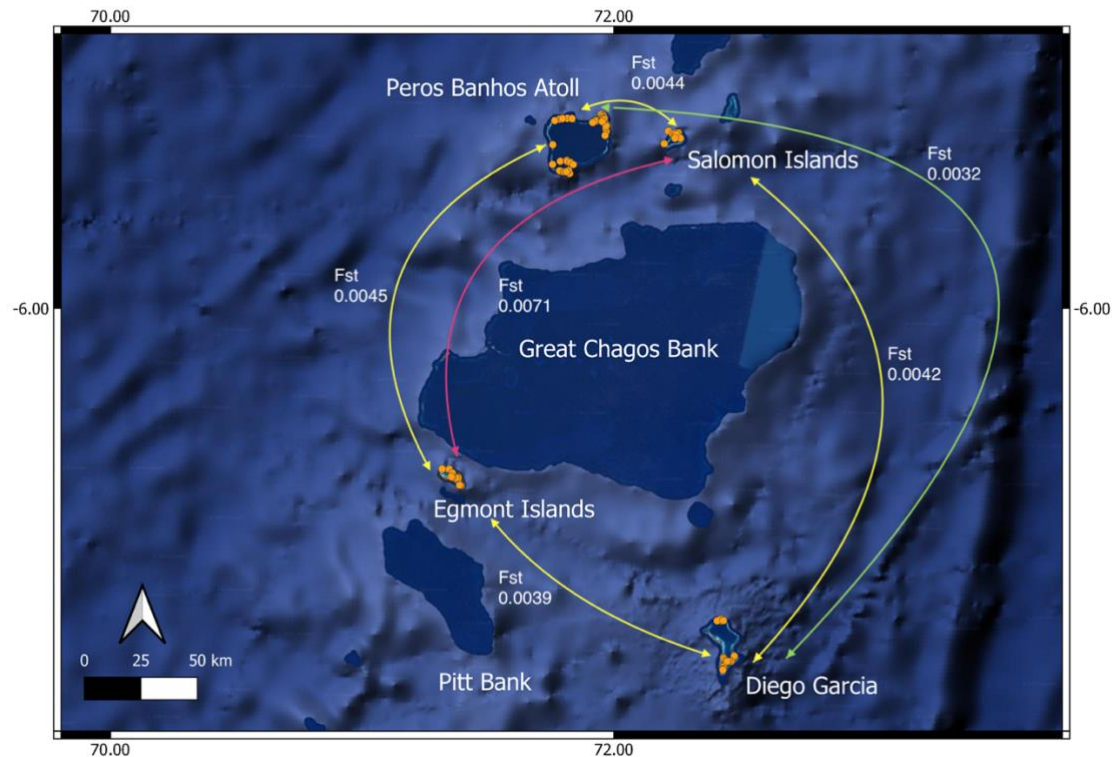


Figure 4.3 *L. bohar* – a representation of population structure between atolls for loci present in >75% of individuals at all sample sites. Red arrow represents highest genetic differentiation; yellow arrow represents moderate differentiation and green arrow represents highest connectivity. Arrows reflect order of differentiation are not indicative of distance travelled or directions of movement of individuals.

4.3.5 Population Structure of *L. gibbus* using pairwise F_{ST} comparison

Lutjanus gibbus demonstrated extremely weak population structure for all pairwise population comparisons between individuals at all atolls when analysing loci present in the majority of individuals (more than 75%) – Table 4.5, Figure 4.7. The greatest differentiation exists primarily between the Northern atolls and central atoll of Egmont (Peros Banhos and Egmont - $F_{ST} = 0.0090$; Salomon and Egmont - $F_{ST} = 0.0079$). By contrast to *L. bohar*, *L. gibbus* also demonstrates population structure between the adjacent Northern atolls of Peros Banhos and Salomon. For *L. gibbus*, the lowest levels of differentiation and hence the greatest genetic connectivity exists between Diego Garcia and Salomon. Similarly, connectivity is also good for populations of *L. gibbus* between Diego Garcia and Peros Banhos, but slightly decreased compared to the genetic connectivity between Diego Garcia and Salomon. (These findings of very weak differentiation between sites are also evident when looking at SNPs present in 80-95% of individuals – for further details refer to Appendix Table A4.7)

Table 4.5 F_{ST} values for pairwise comparisons between all four sampling sites for SNPs present in >75% of individuals for *L. gibbus*. Parameter values for RAD-locus assembly were $M = 5$, $n = 4$ and $m = 3$. Highest level of statistically significant F_{ST} values are in bold. Bonferroni corrected p-value of < 0.05 applied to all values. n is number of individuals.

Species	n	SNPs		Egmont	Salomon	Peros Banhos
<i>L. gibbus</i>	69	11,476	Diego Garcia	0.0066	0.0053	0.0054
			Egmont		0.0079	0.0090
			Salomon			0.0084

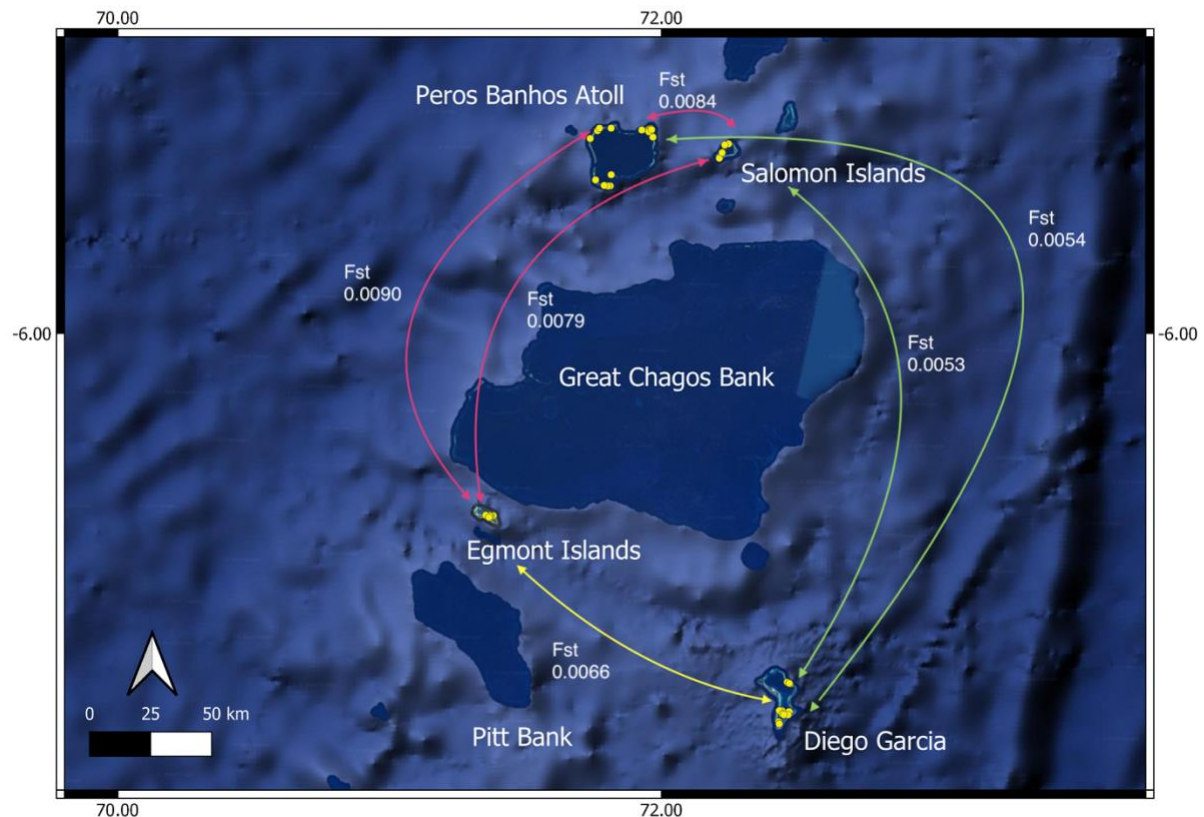


Figure 4.4 *L. gibbus* – a representation of population structure between atolls for loci present in >75% of individuals at all sample sites. Red arrow represents highest genetic differentiation; yellow arrow represents moderate differentiation and green arrow represents highest connectivity. The arrows are not indicative of directions of movement of individuals as this is unknown.

4.3.6 PCA analysis to detect clustering of individuals and populations

To test for the effects of local adaptation on the inference of population structure, structure was analysed using both putative neutral and potential candidate outlier SNP data sets, for each geographical sampling location. Neutral putative SNPs (those at Hardy Weinberg equilibrium (HWE) with a p-value cut-off of < 0.05) were identified using the *populations* analysis tool in Stacks 2.4.1. For *L. bohar*, 20,065 SNPs were used for PCA analysis, and 11,476 SNPs were used for the *L. gibbus* dataset, representing SNPs present in most (more than 75% of individuals) for both species.

Principal component analysis (PCA) performed on each dataset for both neutral, and outlier SNPs (assigned as being potential candidate SNPs under selection, which are not at HWE) did not detect strong clustering defined by geographical location for either species (Figure 4.5 to Figure 4.8). PCA plots for neutral loci in both species suggest admixed populations with a lack of genetic structuring or distinct groups in either *L. bohar* (Figure 4.5) or *L. gibbus* (Figure 4.7). PCA analysis of neutral loci demonstrates a lack of clear differentiation between geographical groups for both species. Focussing on candidate outlier SNPs, for *L. bohar* and *L. gibbus*, there is no clustering defined by geographic locality (Figure 4.6 and Figure 4.8), suggesting complete admixture.

This appears to support the findings from the F_{st} analysis that indicates where genetic structure between atolls exists, it is extremely weak.

(PCA Analysis was also carried out on SNPs present in 80-95% of individuals at all sites, for both species and showed similar patterns of admixture – for further details refer to Appendix Figures A4.8 to A.11)

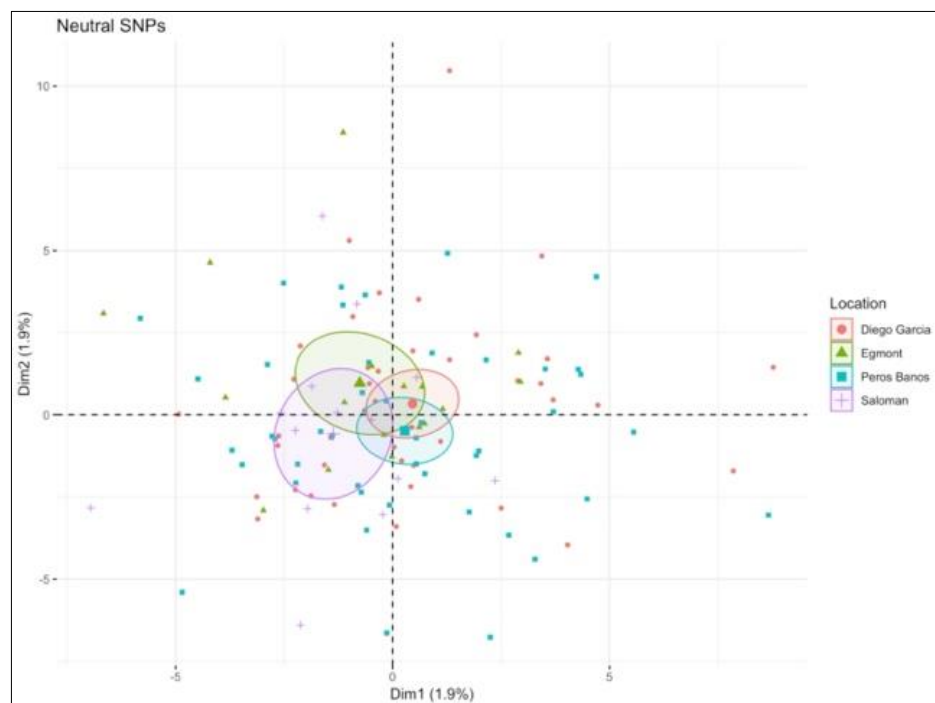


Figure 4.5 Principal component analysis of allele frequencies for neutral loci present in 75% of the *L. bohar* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Ovals represent 95% confidence inertia ellipses.

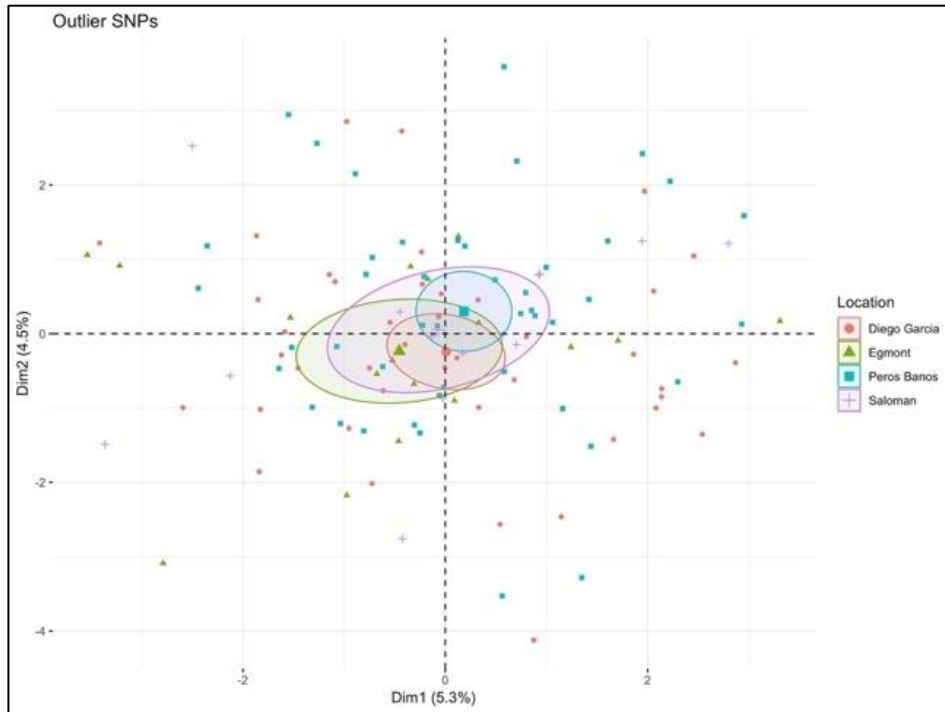


Figure 4.6 Principal component analysis of allele frequencies for outlier loci present in 75% of the *L. bohar* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Ovals represent 95% confidence inertia ellipses.

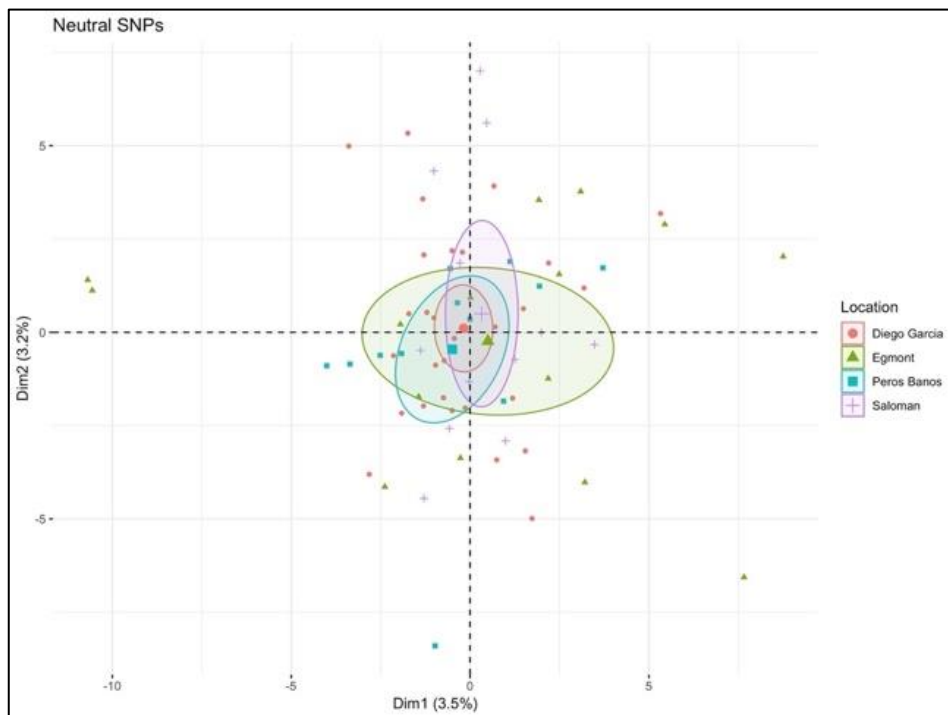


Figure 4.7 Principal component analysis of allele frequencies for neutral loci present in 75% of the *L. gibbus* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Ovals represent 95% confidence inertia ellipses.

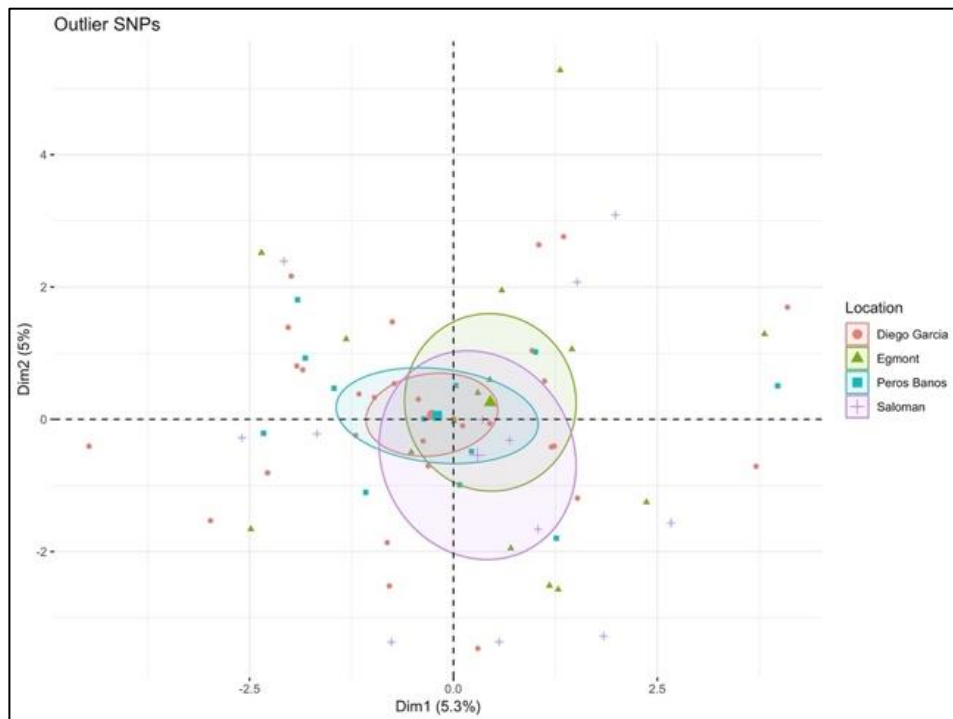


Figure 4.8 Principal component analysis of allele frequencies for outlier loci present in 75% of the *L. gibbus* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Ovals represent 95% confidence inertia ellipses.

4.4 Discussion

4.4.1 Population genetic structure of *L. bohar* and *L. gibbus* in the Chagos Archipelago

Population structure analysis using SNP data and F_{ST} differentiation calculations between atolls detected very weak but significant genetic structure between the Chagos atolls for both species. Although very subtle, this differentiation was most pronounced between the northern atolls (Peros Banhos/ Saloman) and the central atoll of Egmont. For both species, when analysing SNPs present in the majority of individuals (more than 75%) at all sampling locations, these low levels of structure between Saloman and Egmont are similar ($F_{ST} = 0.0071$ for *L. bohar* and $F_{ST} = 0.0079$ for *L. gibbus*). *Lutjanus gibbus* showed the greatest differentiation between Peros Banhos and Egmont atolls ($F_{ST} = 0.0090$), and similar levels of structure between the relatively close, adjacent atolls of Saloman and Peros Banhos ($F_{ST} = 0.084$). For both species, the greatest connectivity and lowest levels of structure exist between Diego Garcia and all other atolls. At all atolls, *L. bohar* shows lower levels of differentiation and greater connectivity than *L. gibbus*.

Whilst differentiation across the study site is weak and F_{ST} values appear small in the context of population structure in a terrestrial landscape, previous studies show such levels of genetic differentiation to be relevant in other marine species. For example, studies investigating population

structure of Atlantic cod (*Gadus morhua*) demonstrated that low levels of genetic differentiation ($F_{ST} = 0.0037$) in marine fish are biologically significant, even at small geographical scales (Knutsen *et al.*, 2011). Other study sites in the Indian Ocean analysing population structure using microsatellite loci, demonstrated low, but significant genetic subdivision in reef fish inhabiting similar sized reef systems as Chagos ($F_{ST}=0.002$; Underwood, Travers and Gilmour, 2012), indicating that low levels of genetic subdivision may occur over relatively small distances. For most species of marine fish, these low levels of population structure are evidence of well connected populations (Underwood, Travers and Gilmour, 2012) and signals of weak differentiation may only be considered a warning signal for heavily exploited populations such as cod and mackerel.

PCA analysis of genotypes for both neutral and potential outlier loci, failed to detect clear structuring between sampling locations. In combination with the F_{ST} analysis of genotypes, this indicates that populations of both species are well mixed and connected in a panmixia. These findings are not surprising, given that gene flow and mixing of individuals between populations at different atolls is primarily mediated via larval dispersal, which potentially enables long-distance mixing of marine organisms (Underwood, Travers and Gilmour, 2012). Additionally, the reefs of the Great Chagos Bank could provide multiple stepping-stone locations for connectivity of individuals between atolls, and opportunities for gene flow between reefs and across the Indian Ocean (Graham, Spalding and Sheppard, 2010; Sheppard *et al.*, 2012).

Other studies have identified weak genetic subdivision, with similar F_{ST} values to our research and also failed to resolve populations into clusters using PCA (Rodriguez-Ezpeleta, Álvarez and Irigoien, 2017). Whilst weak differentiation and low F_{ST} values may be dismissed when taken in combination with other tests (such as PCA) inferring admixture, research has shown that subtle differentiation (identified using F_{ST}) can be biologically significant, with implications for population connectivity and gene-flow in the long-term (Knutsen *et al.*, 2011). Therefore, it is important to explore potential reasons for even subtle genetic differentiation within populations as indicated by F_{ST} , even in the absence of PCA clustering.

The findings of this research suggest that multiple drivers may be responsible for high levels of connectivity for populations of *L. bohar* and *L. gibbus* across the MPA. There is some evidence of weak signals of genetic subdivision operating between the Northern atolls and central region, for sub-populations of both species, but further investigation accounting for local variation within populations would be necessary (e.g. F_{IS} and homozygosity / heterozygosity indices). Potential explanations for the findings are discussed in more detail below.

4.4.2 Factors influencing population structure, gene flow and connectivity between atolls in the Chagos Archipelago

The main process contributing to gene flow of coral reef fish between atolls is dispersal of offspring derived from existing parental genotypes. Gene flow in lutjanids is primarily mediated via pelagic larval dispersal, following seasonal spawning aggregations when tens of thousands of adults gather (Claro *et al.*, 2003; Nanninga and Spaet, 2017). Larval dispersal is influenced by oceanographic factors including tides, currents, temperature, and interactions between ocean fronts and water bodies (Cowen, 2002), as well as physical features such as eddies and upwellings which may restrict the mixing of pelagic larvae (White *et al.*, 2010). Evidence of good connectivity between populations suggests that some or all of these processes that promote mixing of individuals are occurring and thereby facilitating larval dispersal and genetic connectivity (Endler, 1973).

Geographical distance plays a pivotal role in determining the pattern of genetic differentiation spatially within populations of a species, with local dispersal of individuals generating an accumulation of genetic differentiation with increasing distance (Wright 1943, Nei 1973). This genetic pattern of structural distribution, or “Isolation by Distance” will influence the degree of connectivity of sub-populations of lutjanids between the atolls of the Chagos Archipelago. Geographical distribution and distances between atolls will therefore influence dispersal success of pelagic lutjanid larvae and associated gene flow, determining the degree of long-distance connectivity between atolls.

In addition to geographic distance, geophysical factors are also likely to influence the ability of pelagic larvae to disperse between atolls. Many oceanographic processes affecting larval transport, occur in the open ocean and influence dispersal of lutjanid larvae, affecting how far individuals travel and therefore, gene flow. Hydrodynamic drivers influence water movement within lagoons, (such as those found in the Chagos atolls), and include dynamic processes such as tidal flushing, deep-water channels, wind-driven circulation and variability in temperature and salinity. These oceanic features, controlled by tidal forces, affect the exchange of waters between open ocean and lagoons (Pugh and Rayner, 1981; Chevalier *et al.*, 2017; Green, Lowe and Buckley, 2018), potentially influencing larval dispersal and are known to occur in the waters surrounding the atolls of the Chagos Archipelago (Sheehan *et al.*, 2019). Whilst local currents and oceanographic features may promote larval dispersal and connectivity, some oceanic currents act as barriers to larval dispersal and may restrict movement, with fast-moving currents advecting larvae away from neighbouring reefs (Baums *et al* 2006; Wood *et al.* 2016).

Oceanic currents flow past Chagos towards the Western Indian Ocean and East Africa from April to December, switching direction for the rest of the year to flow towards South East Asia (Couper, 1987). It is likely that planktonic larvae, including lutjanid species, are carried along these currents, promoting connectivity between atolls. Indeed, in the context of the wider West Indian Ocean, Chagos may be an important stepping-stone for connectivity with the shallow reefs of the Cocos Keeling Islands to the East (2750 km away) and the Seychelles and Saya de Malas reef systems (1050km away) to the West (Sheppard *et al.*, 2012).

The atolls of Chagos studied in this research are separated by distances of approximately 150-200 km, meaning that pelagic larvae would have to travel these distances to promote gene flow between atolls. Information regarding larval dispersal for lutjanids is scant, but studies indicate that the larvae of some snapper species travel distances up to 28 km, (Harrison *et al.*, 2012; Green *et al.*, 2015), and up to several hundreds of km (Muths *et al.*, 2012). Other studies describe pelagic larval dispersal durations for lutjanids estimated to be between 21 - 38 days (Herrón and Zapata, 2002), and *L. gibbus* specifically, up to 46 (Soeparno *et al.*, 2012) and 62 days (Wilson and McCormick, 1999). Such pelagic larval durations, combined with estimated average current flow of 0.5 ms^{-1} (Bonjean and Lagerloef, 2002) and distances between the Chagos atolls of approximately 25-200 km potentially allow for mixing of individuals and gene flow, and dispersal of larvae to nearby reefs (Green *et al.*, 2015). Additionally, the submerged reefs, distributed along the North and Eastern sides of the Great Chagos Bank, and shallow reefs and islanded atolls along the West could serve as stepping-stone locations for pelagic larvae and so contribute towards the greater connectivity across the Chagos Archipelago, facilitating connectivity over shorter distances and thereby covering the greater distances connecting atolls.

Previous research also shows that larval dispersal potential varies according to several factors, with some years having optimal combinations of biophysical oceanic conditions and high numbers of larvae, such that occasional dispersal between distant atolls is viable (Underwood, Travers and Gilmour, 2012). *Lutjanus bohar* reproduces throughout the year and *L. gibbus* throughout spring and summer; therefore, providing opportunities for sufficient larval abundance and occasional dispersal between the Chagos atolls, when conditions allow. In addition to larval dispersal, connectivity between atolls is also (to a lesser degree) mediated via adult movement, and there is evidence that some lutjanid species migrate large distances, (tens to hundreds of km; Green *et al.*, 2015), most likely influenced by the requirement to find prey and gathering for spawning events.

Given the evidence of good connectivity across the Chagos Archipelago, it appears that larval dispersal occurs over the distances between atolls, and facilitates mixing between parental genotypes from each population. Our findings suggest that pelagic larvae of both *L. bohar* and *L. gibbus* may be carried along the currents between atolls. This dispersal over great distances may be occasional, when environmental and oceanographic conditions are optimal, such that long-distance larval dispersal occurs in sufficient numbers to perpetuate gene flow throughout the Chagos Archipelago. Further work investigating the biology of the larval phase for these two species, combined with investigation of the oceanographic influences and geomorphological features in the Chagos Archipelago, may assist in detecting the underlying physical processes affecting larval dispersal between atolls. In addition to factors that promote genetic connectivity between marine populations, other processes may also operate that contribute to weak genetic structure. These include 'self-recruitment', 'sweepstakes-reproduction' and fish behaviour (avoidance of predation risk), all of which may influence the retention or dispersal of individuals at a site. (Kingsford *et al.*, 2002; Simpson, Meekan and McCauley, 2004; Almany *et al.*, 2007; Gerlach *et al.*, 2007; Green *et al.*, 2015).

Larval dispersal of marine fish is not a passive event, dictated solely by physical and environmental factors. The movement of pelagic larvae is an active process, whereby larvae are able to orientate themselves to best utilise currents, as well as swim faster than oceanic current flow, and most importantly, they can actively navigate towards reef locations (Leis, 2010). This means that although fish larvae may spend a long time developing in the plankton, which potentially facilitates dispersal over long distances, once they are ready to settle on a reef, they can actively direct themselves to a desired location. Active self-recruitment of individuals to their natal reef has been described for many tropical fish species (Almany *et al.*, 2007; Gerlach *et al.*, 2007; Paris, Cherubin and Cowen, 2007; Underwood *et al.*, 2012). Factors such as nutrient availability and reef architecture may be important drivers in the self-recruitment of larvae (Cowen and Sponaugle, 2009). Previous studies have shown that the physical composition and structure of atolls influences patterns and distribution of larval dispersal and recruitment (Underwood *et al.*, 2009; Underwood, Travers and Gilmour, 2012).

The atolls investigated in the Chagos Archipelago are populated with reef sharks, fish, algae and corals (Sheehan *et al.*, 2019), suggesting they are healthy, productive environments. However, there may be differences at individual reefs which may influence larval dispersal and recruitment (for example, Diego Garcia has a reduced fish biomass compared to the other atolls sampled in this study). The plankton assemblage within lagoons in the Chagos Archipelago is an important source of nutrients for

larval stages of tropical fish (Sheehan *et al.*, 2019) and differences in nutrient availability and release associated with this process at individual atolls may influence juvenile self- recruitment.

Another process that may contribute to the retention of individual genotypes at atolls is ‘sweepstakes reproduction’. This stochastic process influences genetic composition of populations of newly recruited pelagic larvae such that successful settlement of larvae is determined by environmental influences on larval developmental stages, and new recruits in a generation are often derived from a few parents (Christie *et al.*, 2010). These factors may act to reduce genetic diversity at local spatial scales and drive genetic differentiation over time. ‘Sweepstakes reproduction’ has been described for other coral reef fish (*Chronis margaritifer*) at remote atolls in the East Indian Ocean (Underwood, Travers and Gilmour, 2012). Where this process occurs, genetic composition and diversity may vary at different sites, with differences occurring at localised scales and with signals of panmixia apparent at one atoll and reduced genetic connectivity at an adjacent reef.

Fish behaviour in response to predation-risk / prey availability may also influence distributions and subsequent mixing of individuals. Indeed, predation risk assessment has been demonstrated for reef fish during the larval phase with larvae tending to settle in lower risk environments (Dixon *et al.*, 2011). It is not clear to what extent predation risk influences larval distribution in the Chagos Archipelago. However, shark predation is a risk factor for adult lutjanids, until fully grown, when they can compete as mesopredators and also escape gape-limited predation (Barley, Meekan and Meeuwig, 2017b). There is evidence of shark depletion (most likely linked to illegal fishing at the northern atolls of Salomon and Peros Banhos, with Salomon having fewer and smaller sharks compared to Peros Banhos, and Egmont having the greatest abundance and diversity of sharks (Meeuwig, personal communication, 2020). Diego Garcia also has depleted shark stocks, most likely due to the impact of recreational fishing and reduced fish biomass on the island. These patterns of shark distribution may influence how attractive and safe atolls are for lutjanids, with the shark-depleted atolls of Peros Banhos and Salomon providing a less risky environment, with good access to prey and reduced competition from sharks, compared to Egmont.

At Peros Banhos, greater numbers of sharks and increased predation risk, compared to Salomon could create a preference for the smaller, more vulnerable lutjanid, *L. gibbus* to stay at the natal reef of Salomon as juveniles and adults. *Lutjanus bohar* is a bigger species than *L. gibbus* and could therefore more easily compete with smaller shark species for prey, as well as escape predation due to exceeding gape limit of sharks. Consequently, residing at Peros Banhos is less risky for *L. bohar* than *L. gibbus*

and might lead to greater movement and connectivity of *L. bohar* between Peros Banhos and Salomon, compared to *L. gibbus*. The influence of predation risk could be a factor when considering movement of adults between atolls, as well as in determining which atolls are 'safest' for juveniles and the smaller species of *L. gibbus* which is more vulnerable to shark predation than the bigger *L. bohar*. Further work would be necessary to determine if safer atolls are associated with increased larval retention and self-recruitment of individuals.

This study has detected very good genetic connectivity between atolls throughout the Chagos Archipelago and the low F_{ST} values indicate mixing of populations. Whilst there are extremely low signals of differentiation for both species, which could hint at the occurrence of larval self-recruitment to natal reefs, or 'sweepstakes reproduction', care must be taken not to over-interpret these F_{ST} values. Without additional information such as F_{IS} (inbreeding coefficient) and heterozygosity measurements for fish at all atolls, it is not possible to determine whether there are reductions in heterogeneity at atolls as a consequence of either of these processes which act to increase genetic differences between populations. If self-recruitment is operating at specific atolls, higher F_{IS} values indicating an increase in inbreeding and a reduction in the heterogeneity of new recruits at these atolls would be expected. Likewise, the effects of sweepstake reproductive success at different atolls in the Chagos Archipelago could potentially lead to reduced genetic diversity, greater within-population relatedness and contribute to differentiation between atolls. To determine whether this is a contributing factor, further investigations would be necessary, comparing adult and recruit samples to identify reduced allelic diversity and increased relatedness within recruits, indicating potential sweepstakes processes taking place. In conclusion, the null hypothesis can be accepted as there is evidence of weak population structure between sites in the Chagos Archipelago indicating that populations are well connected.

5. Concluding Comments

In the face of ongoing climate change and environmental degradation, an understanding of the effects of global warming, over-exploitation and pollution on the world's oceans and marine ecosystems is required. A key goal for scientific research is to develop our knowledge of the response of marine ecosystems and communities to environmental change and disturbance. Understanding the natural variation in marine ecosystems and the ability of marine fish and animals to adapt to environmental change is important for understanding the processes that underpin a sustainable and thriving marine environment.

A changing environment has the potential to influence marine species in a number of ways: Environmental variability occurring over a rapid or intermittent timeframe can cause a change in species distribution; it can induce acclimatisation to the altered environment as a result of phenotypic plasticity (Hoffmann and Sgrò, 2011). Alternatively, species may respond to ongoing shifts in environmental change, that occur over several generations by adaptation of the genotype, via evolutionary selective pressure (Hoffmann and Sgrò, 2011). This can lead to genetic subdivision as differences between populations accumulate over time, resulting in changes to population genetic structure.

Fish are a highly plastic group and their inherent ability to respond to environmental change generates a diversity of variation in fish phenotypes, such that morphological, behavioural, and physiological traits are best suited to the current environmental conditions. This plasticity of response is particularly acute during the early developmental stages of life but is also important in adult stages. For example, metabolic, physiological, locomotive, reproductive and behavioural traits, controlled by endocrine and enzymatic pathways may also be modulated by altered responses to environmental change (Jonsson and Jonsson, 2014).

Fish are particularly exposed to the effects of global change, human exploitation, and the resulting variation in a changing marine environment. And yet, the extent to which most fish species can respond phenotypically is not well studied for many species, and particularly those that are less commercially important. The lack of physical barriers in the marine environment and a pelagic larval phase enables marine organisms to disperse long distances using oceanic currents, thereby connecting distant populations genetically and demographically. Maintaining genetic connectivity between populations is important for conserving genetic heterogeneity of marine fish and is

particularly significant in the context of multiple anthropogenic disturbances and environmental change. Over exploitation and habitat degradation, can potentially leading to local depletions or extinctions of subpopulations of fish (Worm *et al.*, 2006) and a subsequent loss of genetic diversity. A loss of genetic diversity is problematic for a species as it reduces the capacity to adapt to new environmental circumstances (Mace and Purvis, 2008).

The two focal species in this study, *L. bohar* and *L. gibbus*, are large mesopredatory piscivores, that help to shape trophic structure in marine ecosystems (Ruppert *et al.*, 2013; Barley, Meekan and Meeuwig, 2017a). They are heavily exploited throughout their range within the Indo-Pacific Oceans (Marriott, Mapstone and Begg, 2007). Despite the important role they play in marine ecosystems, they are not well studied, and thus the potential effects of anthropogenic pressures (both existing and newly introduced) cannot be anticipated. There is therefore a need to acquire more information regarding the biology, existing population structure and potential response to anthropogenic disturbances and changing environmental pressures, for these species.

It is likely that marine fish are experiencing the effects of environmental change due to anthropogenic pressures, above and beyond what would be considered the environmental norm. These effects are likely to be exerted before we have been able to document baseline measures in terms of morphological variation and population structure. Such 'baseline' measures are crucial if we are going to reliably measure change over time and use this information to best maintain sustainability of marine species.

There are few places left in the world's oceans that are unaffected by anthropogenic impacts. In order to measure baselines of population structure and morphological variation for the two reef associated lutjanid species in this study, it was necessary to carry out the research at one of the few remaining coral reefs that are considered to be as near to 'pristine' as is possible. The study was therefore conducted at The Chagos Archipelago, which due to its remote location in the middle of the Indian Ocean, is distanced from human impacts and most direct anthropogenic pressures (Sheppard *et al.*, 2012).

This study adopted a novel approach to explore morphological and genetic variation for two lutjanid species throughout the Chagos Archipelago. We believe this to be the first study of its kind, carried out in a relatively pristine coral reef environment. Whilst other research has identified morphological variation for lutjanid species, these studies have not applied a genomic approach in parallel (Barley,

Meekan and Meeuwig, 2017a; Barley, Meekan and Meeuwig, 2017b; Hammerschlag *et al.*, 2018). These studies acknowledge that although phenotypic differences in response to altered environments can be identified, they are unable to identify the extent to which morphology is the result of plasticity or evolution via natural selection.

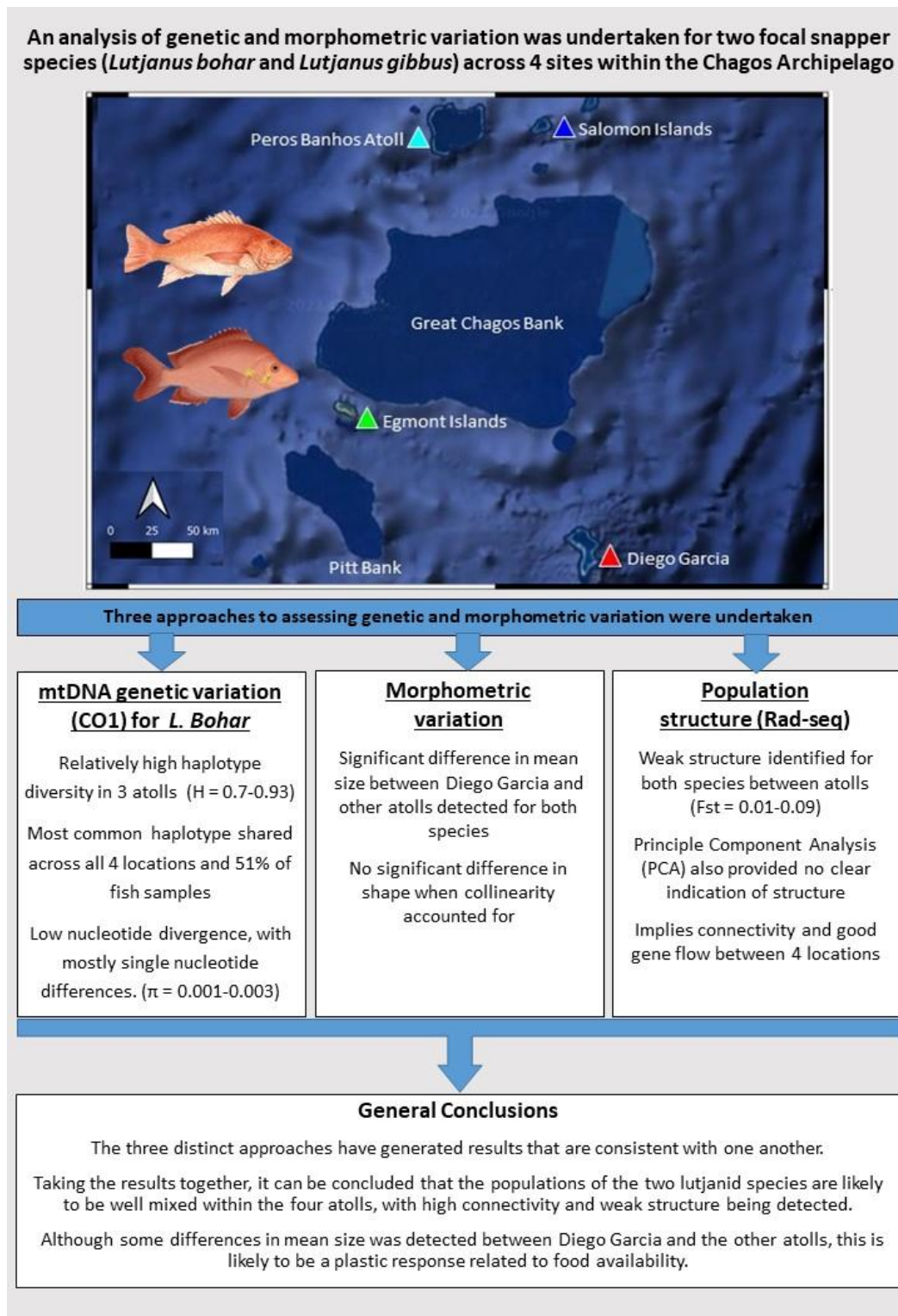


Figure 5.1 – Summary infographic of research

Furthermore, whereas additional studies have investigated morphological variation of a fish species in conjunction with an analysis of genetic variation, to the best of our knowledge, the genetic element of the research has only included the analysis of genetic variation for mitochondrial genes or low numbers of microsatellite data. This study includes a genome-wide analysis of population structure, using SNP data, which has been shown to provide a more accurate inference of population structure than traditional studies using single or a few polymorphic markers (Benestan *et al.*, 2015).

As synthesised in Fig 5.1, this study combined three separate approaches – morphometric analysis, CO1 mitochondrial gene analysis, and RAD-sequencing genome analysis – to assess morphometric variation, genetic variation and population structure for *L. bohar* and *L. gibbus*.

Significant differences in size (as measured by fork length, body depth and circumference) were identified for both species, when comparing populations between atolls in the Chagos Archipelago. This difference in size was almost entirely explained by variations between Diego Garcia and all other atolls. Individuals were smaller at Diego Garcia than conspecifics at the northern atolls and Egmont. We suggest that this morphological variation is the result of a better diet due to higher fish biomass, at Egmont, Salomon and Peros Banhos. We also believe that the primary driver that gives rise to differences in diet and food availability (and hence size), is the recreational fishery at Diego Garcia which appears to have directly contributed to the depleted fish biomass at this atoll. In addition to depleting fish biomass, fishing at Diego Garcia has most likely directly affected lujanids, as surveys show that fish of higher trophic levels (such as the two species in this study) have been directly targeted (Sheppard *et al.*, 2012). Evidence of reduced sizes for *L. bohar* and *L. gibbus* at Diego Garcia suggests that in common with other locations, largest individuals for both these species may have also been fished in preference to smaller fish, at this atoll (Dulvy *et al.*, 2004). It should also be noted that there is evidence of illegal fishing operating within the waters of the Chagos MPA. *L. bohar* and *L. gibbus* are both fished for human consumption, and whilst there is some data regarding illegal catch composition, monitoring is limited, and information is scarce (Sheppard *et al.*, 2012; Graham *et al.*, 2013). The full effects of illegal fishing on the two focal species in this study and the extent to which they are fished in the region are not known.

Other drivers that can contribute to morphological variation were discounted for various reasons. Whilst structural complexity of the reef can drive morphological change (Olsson and Eklöv, 2005; Mihalitsis and Bellwood, 2019) and varies between the atolls within the Chagos Archipelago (Sheppard

et al., 2012; Graham *et al.*, 2013), morphology associated with habitat structural complexity was not detected for the fish in this study. Body depth has been shown to increase in relation to fork length in more complex habitat and this was not detected for *L. bohar* or *L. gibbus* (Olsson and Eklöv, 2005; Mihalitsis and Bellwood, 2019).

Likewise, any effects caused by predation risk differences between the atolls were also ruled out as potential drivers of the differences in size that were observed. Increased predation risk can drive morphological change such that fish are more streamlined and less deep bodied for a given fork length, with bigger caudal fins and larger eyes (traits that enhance vigilance and escape from predators (Hammerschlag *et al.*, 2018)). Although there was data indicating that predation risk is higher at the northern atolls and Egmont than it is at Diego Garcia (Meeuwig, 2020, personal communication), both *L. bohar* and *L. gibbus* did not display morphologies associated with predation risk. Individuals of both species showed no differences in shape between atolls, regardless of predation risk.

The impact of temperature on the two lutjanids in this study was also considered and whether this may have affected fish morphology giving rise to the size differences identified between atolls. Sea surface temperatures increased considerably throughout the Chagos Archipelago during 2015 and 2016 due to global climatic warming events and caused mass coral bleaching. Whilst evidence suggests that temperature increases were more severe in the southwest of the Archipelago (in the region around Egmont) in 2015 and 2016, the thermal stress experienced between all atolls was extreme (Head *et al.*, 2019). Therefore, it is likely that fish at all atolls were similarly affected by the increases in sea surface temperature. Without specific temperature data for each atoll where fish were sampled in our study, it is difficult to draw any conclusions with regards to the effect that sea surface temperature may have had on the lutjanids at individual atolls, and the extent to which this could have contributed to morphological differences between the atolls.

The investigation of genetic variation using mitochondrial CO1 gene analysis indicated that most individuals share the same haplotype for this gene, throughout the archipelago, with any genetic differences being attributable to single nucleotides and only being found in single individuals where they occurred. These findings are in common with the findings from studies investigating haplotype diversity in other lutjanid species (Zhang, Cai and Huang, 2006; Gomes, Sampaio and Schneider, 2012). These findings indicate that *L. bohar* populations are genetically homogeneous throughout the geographical range of the study and well connected across the Chagos Archipelago, most likely as a result of dispersal during the larval stage. The outcomes of neutrality tests do not suggest any recent

or ongoing bottleneck events. We note upfront that analysis of the CO1 gene was only investigated for a single lutjanid species (*L. bohar*) in this study, and similar investigation for *L. gibbus* is recommended to confirm this pattern of genetic variation and haplotype diversity is shared for both species.

RAD-seq analysis provided a genome wide assessment of genetic variation for both species across the Chagos Archipelago. Very weak levels of genetic differentiation were identified for both *L. bohar* and *L. gibbus* between all atolls. This lack of genetic structure is consistent with high levels of connectivity and gene flow between populations of each lutjanid species. Evidence of good connectivity between atolls is not surprising, given that gene flow for these fish is primarily mediated by pelagic larval dispersal (Okiyama, 1988; Drass *et al.*, 2000; Sakaue *et al.*, 2016). The degree of connectivity between populations of lutjanids is influenced by the geographical distance that pelagic larvae must travel. As geographical distances between atolls are relatively small (up to 200 km), these distances are within the potential estimated range for lutjanid pelagic larvae to disperse, using local tidal currents and the South Equatorial Current which flows east-west for part of the year, before reversing for a few months of the year (Obura, 2012). Evidence of well-connected populations of *L. bohar* and *L. gibbus* between the atolls is consistent with other studies for reef dwelling fish in the region (Underwood, Travers and Gilmour, 2012).

Whilst extremely low F_{ST} values were obtained in the RAD-seq analysis, these weak signals of variation were similar between all atolls and are consistent with other studies in genetically well connected marine fish (da Silva *et al.*, 2016; Puritz, Gold and Portnoy, 2016). Weak differentiation could possibly be attributed to low levels of larval self-recruitment to natal reefs (Gerlach *et al.*, 2007; Paris, Cherubin and Cowen, 2007), or the effects of sweepstakes reproductive success (Christie *et al.*, 2010). These processes can contribute to low levels of genetic variation between populations of a species and examples of fine scale partitioning of genomic variation is common in marine fish, even at small spatial scales (Puritz, Gold and Portnoy, 2016). It is not possible within this study to determine whether these processes have contributed to the population genetic variation observed for the two lutjanid species in the Chagos Archipelago. To confirm this, additional experimental and analytical data would be necessary – for example, the inbreeding coefficient (F_{IS}), and observed and expected heterozygosity indices for each sampling site would provide information regarding how closely related individuals are at each atoll. If individuals are closely related or have lower than expected heterozygosity value, this could indicate that larvae are returning to the reef they originated from, rather than mixing with populations at distant reefs. To determine whether sweepstakes recruitment has affected genetic

variation at atolls, the degree of inbreeding and observed and expected heterozygosity in parents would need to be compared to juveniles. If only a few parental genotypes have passed to the next generation, a higher degree of inbreeding and reduced heterozygosity would be expected in juveniles, than in parents.

The findings of these three approaches would appear to be consistent with the hypothesis that there is little population structure for *L. bohar* and *L. gibbus* within the Chagos Archipelago and that populations are genetically well connected. Haplotype network analysis of CO1 variation demonstrates that for *L. bohar* at least, haplotypes are shared across the entire Archipelago, and this indication of connectivity is also supported by the lack of population structure detected for both species in the genomic analysis using SNP data. Similarly, when all findings are looked at in combination, the results from this study imply that although there are morphological variations for populations of both species between Diego Garcia and the atolls of Egmont, Salomon and Peros Banhos, the differences in size are not the result of genetic differences resulting from adaptive evolution. Rather, these differences in size are likely to represent plasticity of phenotype that is driven in response to differences in food availability and nutritional value of diet, between Diego Garcia and the other atolls.

In reaching these conclusions, an awareness of the limitations of this study needs to be considered. When carrying out a study such as this one, in a natural environment, it is difficult to control for the multiple factors that could be exerting an effect on the ecosystem. These limitations could be broadly addressed by collecting additional data for the two fish species, and more data measuring the environmental factors that may contribute to morphological variation.

Additional fish data for both species would help build up a more comprehensive picture of the fish assemblages at each atoll in the Archipelago. This data would ideally include surveys of fish abundance and distribution to determine whether *L. bohar* and *L. gibbus* were less abundant, as well as being smaller at Diego Garcia and how distribution varied for each atoll. Information regarding the age of the fish would also be informative as it could be used to identify juveniles and parents to investigate processes such as sweepstakes reproduction that may occur at specific atolls. However, given the protected status of the region, it is not likely that otolith samples (which can be used for aging fish) could be taken.

An additional limitation with the fish data for this study is the small and unequal sample sizes at each atoll. There were several errors in the recording of the raw morphometric data at the time of sampling, multiple samples did not have complete data records, and these samples could not be used for the morphological analysis. Ideally, more samples would have been collected for each species, at every atoll, to provide a more accurate representation of the lutjanid populations. Another potential limiting factor is that fish were targeted by professional fishermen, and very small individuals were not sampled, due to the permitted licence conditions. This should have only eliminated juveniles from sampling, but it is possible that some small adults were also excluded from the study, and this could skew the size data. However, this method was applied to all sites, so the impact of any missing small adults should be equal across all sites.

It is recommended that CO1 barcoding and haplotype analysis is carried out for *L. gibbus* as to provide information regarding haplotype distribution, population homogeneity and within species variation for the CO1 gene. Unfortunately, it was not logistically possible to carry out CO1 barcoding for both species in this project.

Whilst other studies provided some limited data regarding environmental factors that could vary between sampling sites (Graham *et al.*, 2013; Head, *et al.*, 2019), more comprehensive data recordings of abiotic and biotic factors that can influence fish morphology and physiology would be informative. Survey data measuring environmental drivers such as temperature and reef complexity at each atoll, as well as fish surveys to measure predator abundance and fish assemblage could be used to try and identify whether each of these factors influence fish morphology. In addition to morphological variation in fish, environmental changes (such as sea temperature increases due to climate change) can drive phenotypic plasticity for other behavioural and physiological traits of interest. To investigate these processes, transcriptome analysis could be used to detect differences in expression of endocrine factors that control the physiological processes associated with cognition, feeding, movement and reproduction. This would provide a better understanding of these processes and the limits within which they can operate, to inform conservation decisions in the light of the effect of climate change on the marine environment and the fish that inhabit coral reefs.

It is worth noting that the findings from this study only provide a snapshot in time of morphological and genetic variation in the Chagos Archipelago, as samples were collected on two expeditions that were only months apart. There is no baseline data from previous time points to determine whether these size differences have always existed, or whether they reflect a recently altered environment. As environmental changes operate over multigenerational timescales, a longitudinal study would be

more informative than the snapshot this study provides. However, the limited data from this study could be used as a benchmark for future studies.

In summary, the results of this study confirm that populations of *L. bohar* and *L. gibbus* are genetically well connected between atolls of the Chagos Archipelago. The very low levels of differentiation between atolls are not surprising and are typical of marine species due to long distance larval dispersal and a lack of physical barriers in the oceanic environment. Differences in size between the atolls was attributed to fishing related depletion of larger individuals and reduced biomass and food availability at the only atoll where fishing is permitted (Diego Garcia). These findings provide baseline data regarding population structure for both species and indicate that in a remote reef location and in the absence of intensive exploitation and overfishing, these fish maintain a genetically connected population. Evidence of size differences between atolls suggest that even in a relatively 'pristine' reef environment, the remote location and protected status does not fully safeguard fish against even low levels of fishing related exploitation. The longer-term resilience of these two fish species to ongoing exploitation and environmental change is not certain, as the isolated location and lack of anthropogenic pressures compared to other reef locations cannot fully protect against the global impact of climate warming. Whilst the outlook for these fish within the Chagos Archipelago and MPA is optimistic in the short term, their long-term success will depend on their ability to acclimatise to the additional pressures arising from increasingly frequent and severe consequences of global climate change.

Appendix: Chapter 3

Post hoc Tukey tests to test for differences in size of *L. bohar* between sites

Tukey's HSD Test for multiple comparisons of mean value of weight was significantly different between Egmont and Diego Garcia ($p = 0.035$, 95% C.I. = 0.0130, 0.4926), and between Egmont and Peros Banhos ($p = 0.027$, 95% C.I. = -0.4137, -0.0173), (Table A3.1)

Table A3.1 Tukey HSD Test to compare differences in weight between sites for *L. bohar*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	p adj
Egmont-Diego Garcia	0.25278945	0.01299773	0.4925812	0.0346393
Peros Banhos-Diego Garcia	0.03728298	-0.15354865	0.2281146	0.9566500
Salomon-Diego Garcia	0.23430451	-0.03975655	0.5083656	0.1215242
Peros Banhos-Egmont	-0.21550647	-0.41372785	-0.0172851	0.0274570
Salomon-Egmont	-0.01848495	-0.29774194	0.2607720	0.9981601
Salomon-Peros Banhos	0.19702153	-0.04151801	0.4355611	0.1426461

Tukey's HSD Test for multiple comparisons found that the mean value of circumference was significantly different between Diego Garcia and Egmont ($p = 0.0000176$, 95% C.I. = 0.06936061, 0.226941934), Diego Garcia and Peros Banhos ($p = 0.0153629$, 95% C.I. = 0.01006194, 0.131867738), Diego Garcia and Salomon ($p = 0.0060254$, 95% C.I. = 0.02682403, 0.216762164) and Peros Banhos and Egmont ($p = 0.0264924$, 95% C.I. = -0.14788021, -0.006492664), (Table A3.2, Figure 2.3).

Table A3.2 Tukey HSD Test to compare differences in circumference between sites for *L. bohar*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	P adj
Egmont-Diego Garcia	0.14815127	0.06936061	0.226941934	0.0000176
Peros Banhos-Diego Garcia	0.07096484	0.01006194	0.131867738	0.0153629
Salomon-Diego Garcia	0.12179310	0.02682403	0.216762164	0.0060254
Peros Banhos-Egmont	-0.07718644	-0.14788021	-0.006492664	0.0264924
Salomon-Egmont	-0.02635818	-0.12788438	0.075168026	0.9059693
Salomon-Peros Banhos	0.05082826	-0.03753896	0.139195478	0.4418445

Tukey's HSD Test for multiple comparisons found that the mean value of body depth for *L. bohar* was significantly different between Diego Garcia and Egmont ($p = 0.0000404$, 95% C.I. 0.06382445, 0.222720606), between Diego Garcia and Peros Banhos ($p = 0.0041303$, 95% C. I. 0.02003495, 0.142533444) and between Diego Garcia and Salomon ($p = 0.0018775$, 95% C.I. 0.135670158, 0.03990868), (Table A3.3, Figure 3.4).

Table A3.3 Tukey HSD Test to compare differences in body depth between sites for *L. bohar*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	P adj
Egmont-Diego Garcia	0.143272528	0.06382445	0.222720606	0.0000404
Peros Banhos-Diego Garcia	0.081284199	0.02003495	0.142533444	0.0041303
Salomon-Diego Garcia	0.135670158	0.03990868	0.231431633	0.0018775
Peros Banhos-Egmont	-0.061988329	-0.13313260	0.009155941	0.1109691
Salomon-Egmont	-0.007602369	-0.10997569	0.094770953	0.9974246
Salomon-Peros Banhos	0.054385959	-0.03460713	0.143379049	0.3872795

Tukey's HSD Test for multiple comparisons found that the mean value of fork length for *L. bohar* was significantly different between Diego Garcia and Egmont ($p = 0.0000729$, 95% C.I. 0.06254887, 0.22970787), between Diego Garcia and Peros Banhos ($p = 0.0074399$, 95% C.I., 0.01630670, 0.14327308) and between Diego Garcia and Salomon ($p = 0.0028439$, 95% C.I. 0.03675031, 0.23525868), (Table A3.4, Figure 3.5).

Table A3.4 Tukey HSD Test to compare differences in fork length between sites for *L. bohar*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	P adj
Egmont-Diego Garcia	0.14612837	0.06254887	0.22970787	0.0000729
Peros Banhos-Diego Garcia	0.07978989	0.01630670	0.14327308	0.0074399
Salomon-Diego Garcia	0.13600450	0.03675031	0.23525868	0.0028439
Peros Banhos-Egmont	-0.06633848	-0.14145280	0.00877584	0.1035816
Salomon-Egmont	-0.01012387	-0.11719132	0.09694358	0.9947286
Salomon-Peros Banhos	0.05621461	-0.03602433	0.14845354	0.3896964

Post hoc Tukey tests to test for differences in size of *L. gibbus* between sites

Tukey's HSD Test for multiple comparisons found that the mean value of weight was not significantly different between sites (Table A3.5, Figure 3.6)

Table A3.5 Tukey HSD Test to compare differences in weight between sites for *L. gibbus*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	p adj
Egmont-Diego Garcia	0.08860178	-0.032386963	0.2095905	0.2279301
Peros Banhos-Diego Garcia	0.11222856	-0.004269397	0.2287265	0.0632746
Salomon-Diego Garcia	0.13795088	-0.027807411	0.3037092	0.1369151
Peros Banhos-Egmont	0.02362678	-0.105634767	0.1528883	0.9635643
Salomon-Egmont	0.04934910	-0.125615423	0.2243136	0.8810812
Salomon-Peros Banhos	0.02572233	-0.146167413	0.1976121	0.9794213

Tukey's HSD Test for multiple comparisons found that the mean value of circumference was significantly different between Diego Garcia and Egmont ($p = 0.0419233$, 95% C.I. = 0.001414594, 0.1081814), Diego Garcia and Peros Banhos ($p = 0.0000001$, 95% C.I. = 0.068949034, 0.1716587), Diego Garcia and Salomon ($p = 0.0000791$, 95% C.I. = 0.057040851, 0.2082254) and Peros Banhos and Egmont ($p = 0.0230720$, 95% C.I. = 0.006596328, 0.1244154), (Table A3.6, Figure 3.7).

Table A3.6 Tukey HSD Test to compare differences in circumference between sites for *L. gibbus*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	P adj
Egmont-Diego Garcia	0.05479800	0.001414594	0.1081814	0.0419233
Peros Banhos-Diego Garcia	0.12030386	0.068949034	0.1716587	0.0000001
Salomon-Diego Garcia	0.13263311	0.057040851	0.2082254	0.0000791
Peros Banhos-Egmont	0.06550586	0.006596328	0.1244154	0.0230720
Salomon-Egmont	0.07783511	-0.003079688	0.1587499	0.0639045
Salomon-Peros Banhos	0.01232925	-0.067261804	0.0919203	0.9774177

Tukey's HSD Test for multiple comparisons found that the mean value of body depth was significantly different between Diego Garcia and Egmont ($p = 0.0054223$, 95% C.I. = 0.03842823, 0.2957197), Diego

Garcia and Peros Banhos ($p = 0.0008135$, 95% C.I. 0.06304823, 0.3078260) and Diego Garcia and Salomon ($p = 0.0013805$, 95% C.I. 0.08310988, 0.4466590), (Table A3.7, Figure 3.8).

Table A3.7 Tukey HSD Test to compare differences in body depth between sites for *L. gibbus*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	P adj
Egmont-Diego Garcia	0.16707397	0.03842823	0.2957197	0.0054223
Peros Banhos-Diego Garcia	0.18543709	0.06304823	0.3078260	0.0008135
Salomon-Diego Garcia	0.26488443	0.08310988	0.4466590	0.0013805
Peros Banhos-Egmont	0.01836313	-0.12175698	0.1584832	0.9860806
Salomon-Egmont	0.09781047	-0.09634565	0.2919666	0.5543518
Salomon-Peros Banhos	0.07944734	-0.11062082	0.2695155	0.6947212

Tukey's HSD Test for multiple comparisons found that the mean value of body depth was significantly different between Diego Garcia and Egmont ($p = 0.0059212$, 95% C.I. =0.015249907, 0.1212648), Diego Garcia and Peros Banhos ($p = 0.0000005$, 95% C.I. 0.061131491, 0.1619003) and Diego Garcia and Salomon ($p = 0.0000254$, 95% C.I. 0.064873351, 0.2153051), (Table A3.8, Figure 3.9).

Table A3.8 Tukey HSD Test to compare differences in fork length between sites for *L. gibbus*. Where differences are statistically significant, p value is highlighted purple.

Location	Diff	Lwr	Upr	P adj
Egmont-Diego Garcia	0.06825738	0.015249907	0.1212648	0.0059212
Peros Banhos-Diego Garcia	0.11151590	0.061131491	0.1619003	0.0000005
Salomon-Diego Garcia	0.14008921	0.064873351	0.2153051	0.0000254
Peros Banhos-Egmont	0.04325852	-0.014965531	0.1014826	0.2175300
Salomon-Egmont	0.07183184	-0.008845779	0.1525095	0.0988868
Salomon-Peros Banhos	0.02857331	-0.050405636	0.1075523	0.7803436

Appendix: Chapter 4

Stacks packages and parameters

Table A4.1. Elements of *Stacks* RAD-seq data analysis pipeline used and description of primary functions (Catchen *et al.*, 2011, 2013).

Wrapper	Component	Description
	process_radtags	Cleans, filters and demultiplexes raw reads in fastq format. Discards reads with phred score < 10 (e.g., less than 90% probability of being correct). Allows for some sequencing error, whilst removing degraded reads.
denovo_map.pl	Ustacks	Aligns short-read sequences from process_radtags into matching putative alleles (stacks). Compares stacks to build putative loci <i>denovo</i> for each sample and detects SNPs at each locus using maximum likelihood framework.
	Cstacks	Assembles a catalogue of loci across the population representing homologous loci across all population samples, with loci being matched up by sequence similarity.
	Sstacks	Matches each sample against the catalog
	tsv2bam	Converts data from being ordered by sample, to being ordered by RAD locus, with paired-end reads assigned to corresponding single-end reads.
	Gstacks	Contigs are assembled from paired end reads and overlapped with corresponding single-end locus, with SNPs called for each locus.
	Populations	Calculates population level statistics such as F_{ST} , F_{IS} and π .

Table A4.2. Parameter descriptions, default values defined in *Stacks* v 2.0/2.1. The default values and the values used for the analysis in this research are detailed.

Parameter	Stacks module	Description	Default Value	Values used for this study
M	Ustacks	Minimum number of raw reads to form a putative allele (stack)	3	3
M	ustacks	Number of mismatches allowed between putative alleles (stacks) during merging of stacks into a putative locus	2	5
N	cstacks	Number of mismatches allowed between putative loci (stacks) during catalog construction	1	4

Parameter choice selection for loci assembly: *denovomap.pl*

Distribution of the number of SNPs per locus for each iteration of M/n was plotted. Increasing the value of the M and n parameters affected the number of assembled loci and polymorphic loci shared by 80% of samples or more for both datasets (Table A4.3, Table A4.4). Distribution of the number of SNPs per locus was also affected by increasing values for M/n (Figures A4.1 to A4.4).

Setting the parameter value for M (distance allowed between stacks) for *L. bohar*

For the *L. bohar* r80 loci data set, the number of loci shared by 80% of samples or more begins to plateau between M = 4 and M =5 (Table A4.3, Figure A4.3), with a value of M = 5 being sufficient to stabilise the proportion of loci with 1 to 5 SNPs (Figure A4.4). An optimal value of M=5 was selected for stack assembly.

Table A4.3 Increasing the value of M and n and the effect on the number of loci and polymorphic loci that are assembled and shared by 80% or more of samples for the *L. bohar* dataset

M/n value	Assembled Loci/r80	Polymorphic Loci/r80	Variant Sites/SNPs	Polymorphic Sites/SNPs
1	46698	44314	158139	158139
2	48790	46952	196343	196343
3	49142	47584	218504	218504
4	49134	47772	235325	235325
5	48942	47658	247244	247244
6	48790	47518	256116	256116
7	48577	47345	262236	262236

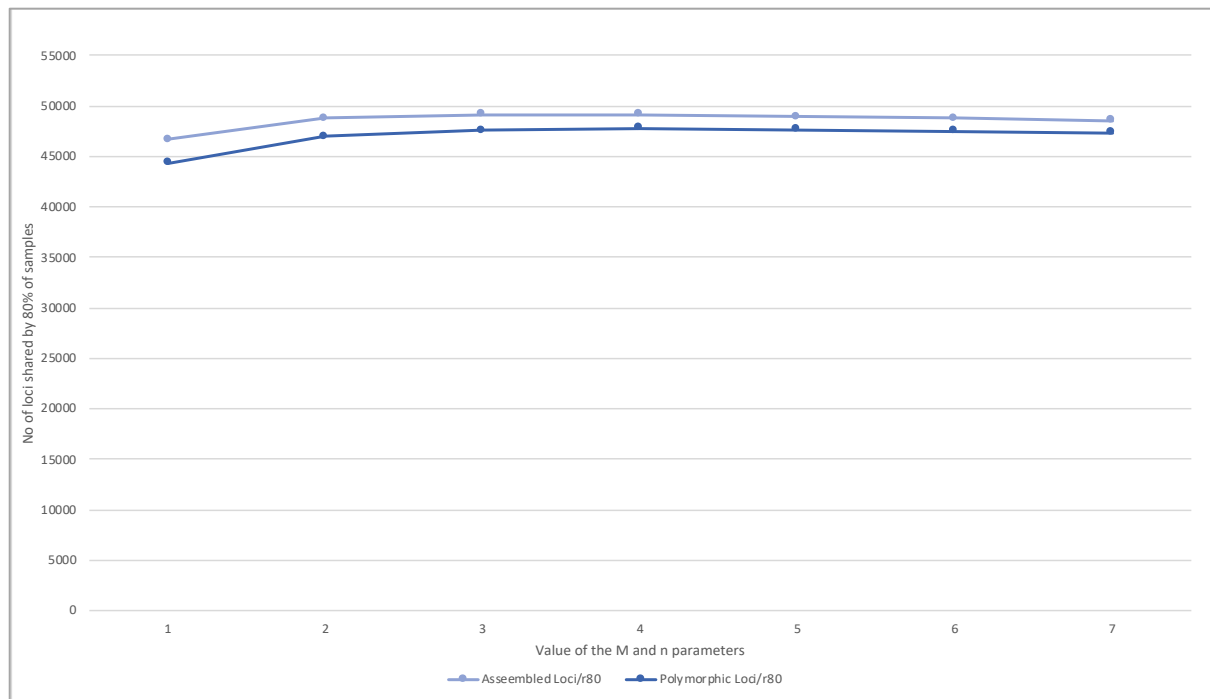


Figure A4.1. *L. bohar* r80 loci - the effect of increasing the value of M and n on the number of loci and polymorphic loci assembled and shared by 80% of samples or more. The M and n parameters are kept equal, and the m parameter is fixed at a value of 3.

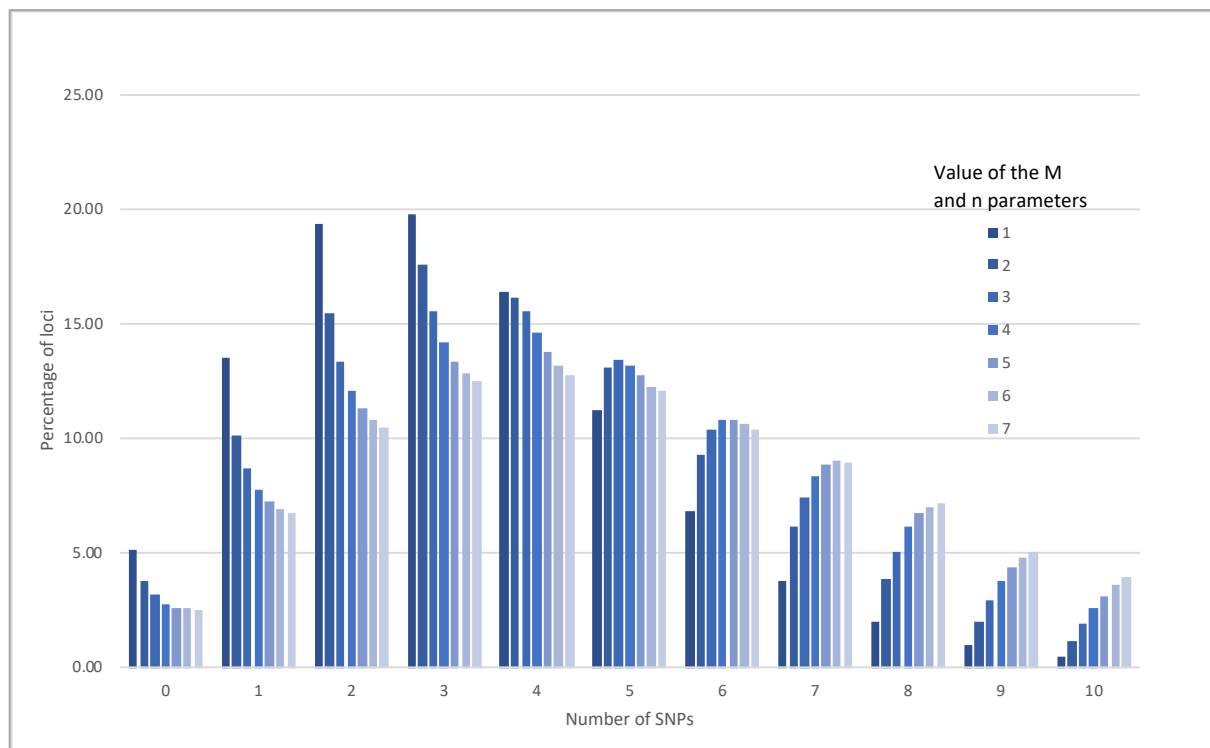


Figure A4.2 Distribution of the number of SNPs per locus with increasing values for M and n for the *L. bohar* r80 loci dataset. The M and n parameters are kept equal, m parameter is fixed at a value of 3.

Setting the parameter value of M (distance allowed between stacks) for *L. gibbus*

For the *L. gibbus* r80 loci data set, the number of widely shared loci begins to plateau starting at M = 5 (Table A4.4, Figure A4.5) and M = 5 is an adequate value to stabilise the proportion of loci with 1 to 5 SNPs (Figure A4.6). A value of M = 5 was selected as the optimal value for the *L. bohar* data set.

Table A4.4 Increasing the value of M and n and the effect on the number of loci and polymorphic loci that are assembled and shared by 80% or more of samples for the *L. gibbus* dataset

M/n value	Assembled Loci/r80	Polymorphic Loci/r80	Variant Sites/SNPs	Polymorphic Sites/SNPs
1	37051	29817	70036	70036
2	40537	34797	102836	102836
3	41848	37035	127501	127501
4	42467	38239	148239	148239
5	42695	38838	164367	164367
6	42722	39090	177376	177376
7	42680	39231	188375	188375

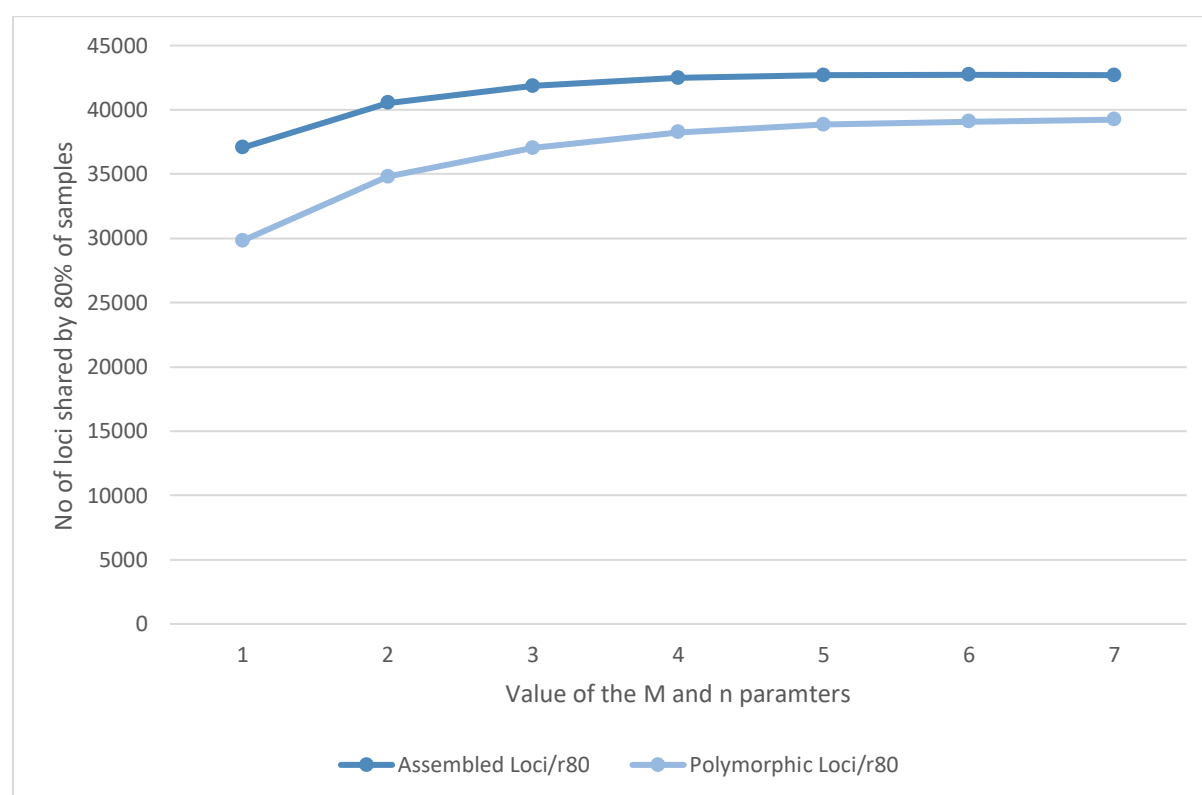


Figure A4.3 Selecting the assembly parameters for *de novo* loci assembly and analysis for the *L. gibbus* r80 loci dataset. The effect of increasing the value of M and n on the number of loci and polymorphic loci assembled and shared by 80% of samples or more. M and n parameters are kept equal, and m parameter is fixed at 3.

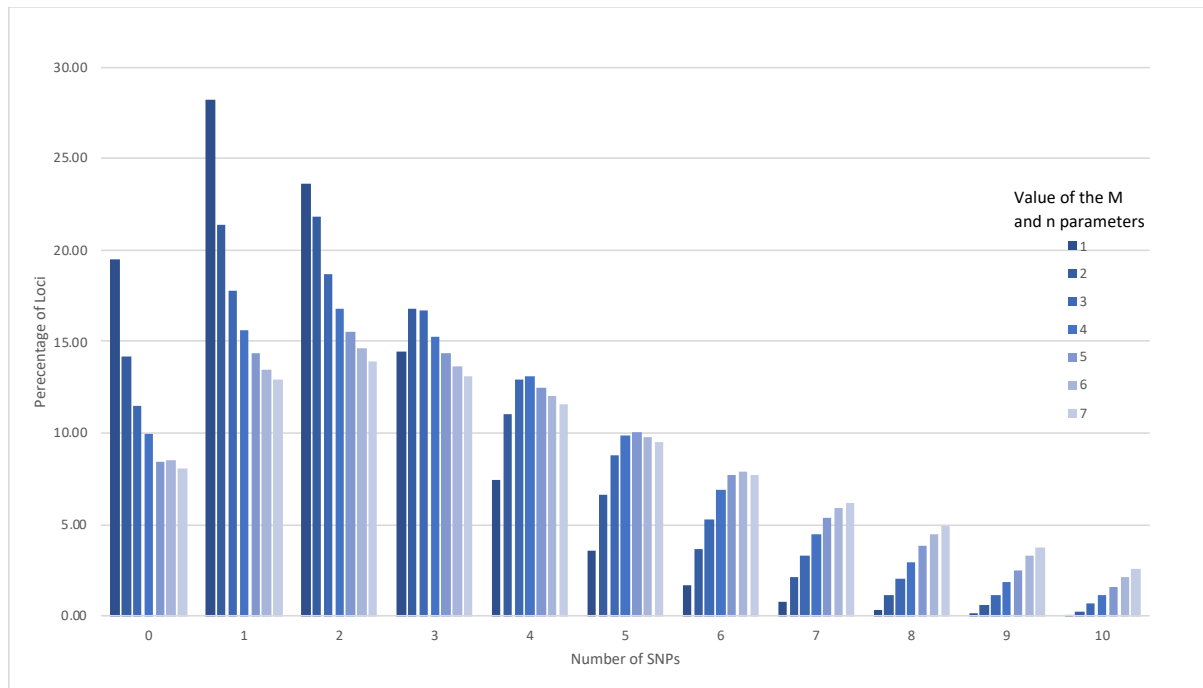


Figure A4.4 Distribution of the number of SNPs per locus with increasing values for M and n for the *L. gibbus* r80 loci dataset. The M and n parameters are kept equal, and the m parameter is fixed at a value of 3.

Setting the parameter value for n (mismatches allowed between catalog loci)

Having identified the optimal value for M at 5 for both data sets, in order to determine the best value to use for the parameter n iterations of n of 4 ($n = M - 1$), 5 ($n = M$) and 6 ($n = M + 1$) were run using the *denovomap.pl* wrapper. Numbers of assembled loci, polymorphic loci and SNPs were calculated for the *r80 loci* for each output data set (Table A4.5). A value of $n = 4$ gave the highest number of polymorphic loci for both datasets.

Table A4.5. Iterating the value of n and effect on the number of assembled loci and polymorphic loci shared across 80% of the population for $M = 5$. Values of n that gave the highest numbers of assembled and polymorphic loci are in bold.

M = 5, populations at 80% (r80 loci)	<i>L. bohar</i>			<i>L. gibbus</i>		
	n = 4	n = 5	n = 6	n = 4	n = 5	n = 6
Number of Assembled Loci	49134	48963	49015	42855	42688	42449
Number of Polymorphic Loci	47772	47672	47773	39220	38825	38248
Number of SNPs	235325	247377	257248	177603	164336	149560

Table A4.6 Genotyping datasets used for population analysis. Shows number of individuals (n), min, max, and mean numbers of loci per sample (Min loci, Max loci, Mean loci), min, max and mean numbers of reads per sample (Min reads, Max reads, Mean reads), min, max and mean coverage per sample (Min coverage, Max coverage, Mean coverage), number of assembled loci, number of polymorphic loci and total number of SNPs.

	<i>L. bohar</i>	<i>L. gibbus</i>
N	118	69
Min loci	20,939	31,583
Max loci	83,267	73,449
Mean loci	61,312	61,886
Min reads	363,745	644,641
Max reads	8,238,457	14,000,000
Mean reads	2,800,560	3,600,000
Min coverage	17 x	20 x
Max coverage	100 x	199 x
Mean coverage	44 x	53 x
Assembled loci	337,131	313, 438
Polymorphic loci	112,216	112,738
SNPs	2, 230,000	1, 790,000

Table A4.7 F_{ST} values for pairwise comparisons between all four sampling sites for *L. gibbus* and *L. bohar*. Parameter values for RAD-locus assembly were $M = 5$, $n = 4$ and $m = 3$. Highest level of statistically significant F_{ST} values are in bold. Bonferroni corrected p-value of < 0.05 applied to all values.

% <i>L. bohar</i> Individuals with locus	No. of SNPs		EG	SA	PB	% <i>L. gibbus</i> Individuals with locus	No. of SNPs		EG	SA	PB
75	20065	DG	0.0039	0.0042	0.0032	75	11476	DG	0.0066	0.0053	0.0054
		EG		0.0071	0.0045			EG		0.0079	0.0090
		SA			0.0044			SA			0.0084
80	13016	DG	0.0037	0.0038	0.0030	80	6396	DG	0.0059	0.0050	0.0051
		EG		0.0063	0.0041			EG		0.0074	0.0082
		SA			0.0041			SA			0.0082
85	7056	DG	0.0034	0.0036	0.0027	85	1177	DG	0.0059	0.0048	0.0048
		EG		0.0057	0.0035			EG		0.0056	0.0078
		SA			0.0036			SA			0.0080
90	1626	DG	0.0033	0.0029	0.0027	90	428	DG	0.0044	0.0028	0.0045
		EG		0.0062	0.0036			EG		0.0033	0.0066
		SA			0.0037			SA			0.0069
95	55	DG	0.0000	0.0016	0.0014	95	2	DG	0.0000	0.0000	0.0000
		EG		0.0045	0.0023			EG		0.0000	0.0000
		SA			0.0011			SA			0.0000

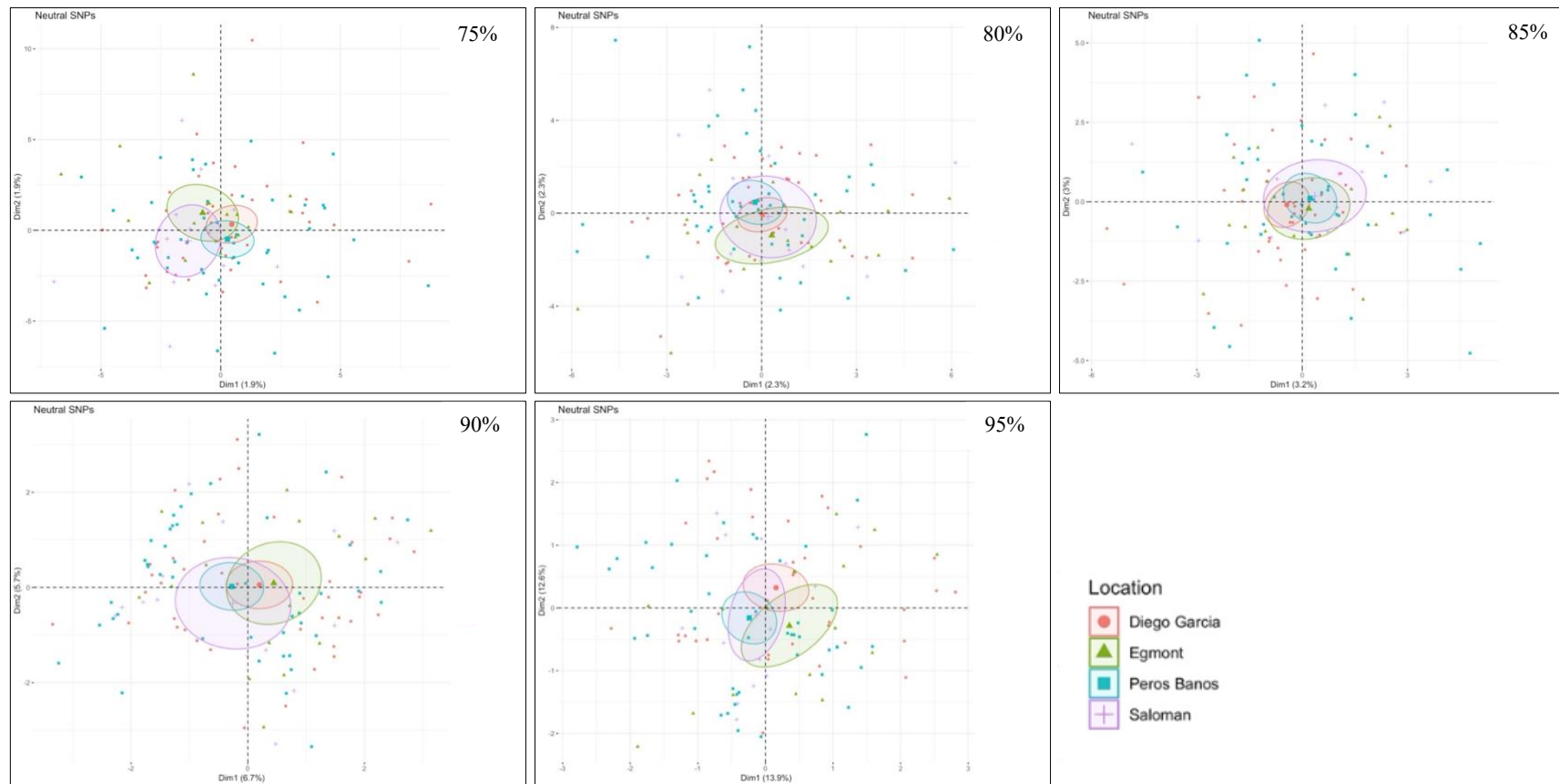


Figure A4.8 Principal component analysis of allele frequencies for neutral loci present in 75,80,85,90 and 95% of the *L. bohar* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Percentage of individuals with loci is detailed at top right of each plot. Ovals represent 95% confidence inertia ellipses.

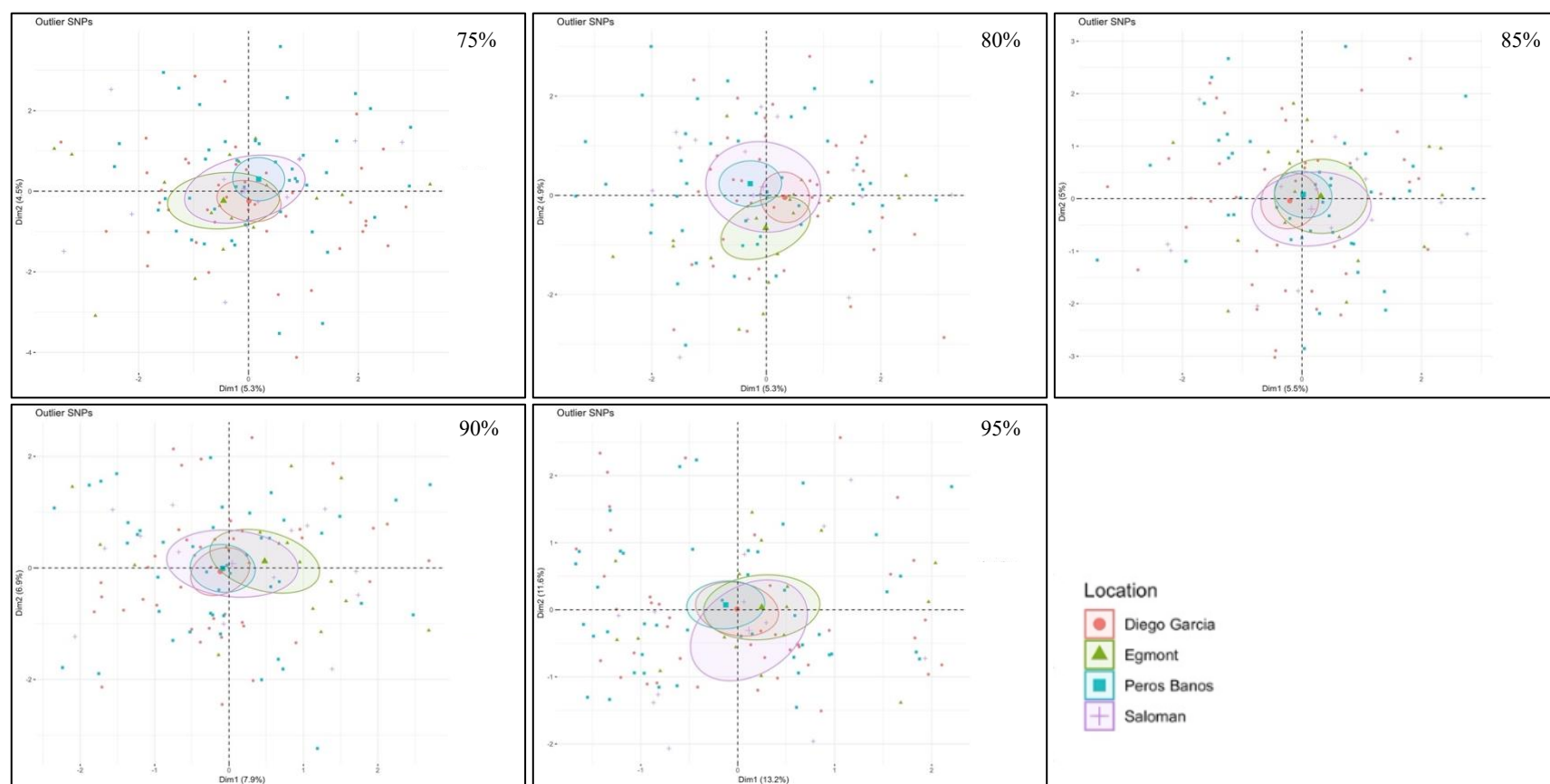


Figure A4.9 Principal component analysis of allele frequencies for outlier loci present in 75,80,85,90 and 95% of the *L. bohar* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Percentage of individuals with loci is detailed at top right of each plot. Ovals represent 95% confidence inertia ellipses.

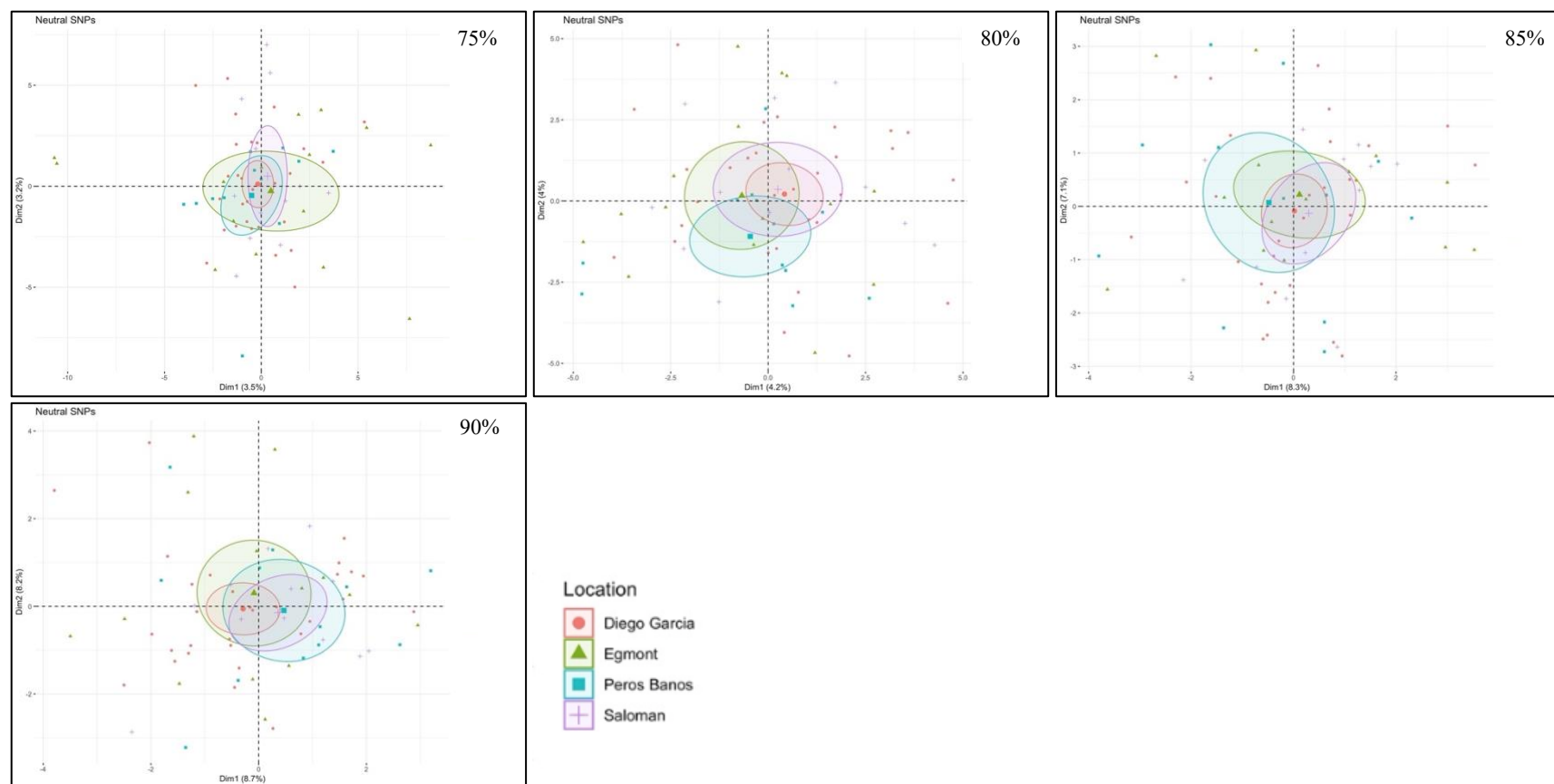


Figure A4.10 Principal component analysis of allele frequencies for neutral loci present in 75,80,85 and 90% of the *L. gibbus* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Percentage of individuals with loci is detailed in box at top right of each plot. Ovals represent 95% confidence inertia ellipses

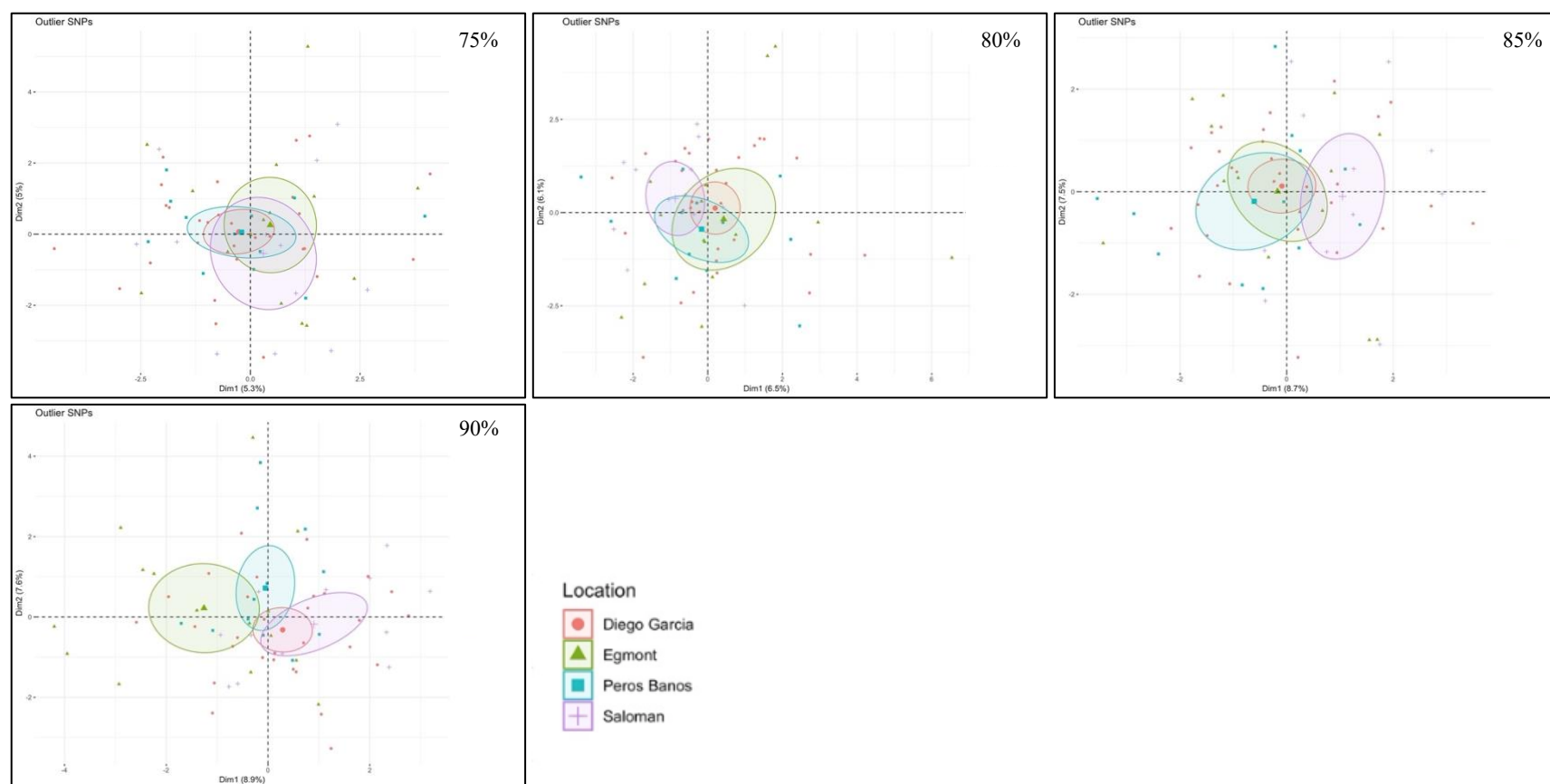


Figure A4.11. Principal component analysis of allele frequencies for outlier loci present in 75,80,85 and 90% of the *L. gibbus* population. Each sample is represented by a coloured dot corresponding to sampling location. Each axis is labelled with the contribution of each principal component (Dim1 and Dim2) explaining variance in the data. Percentage of individuals with loci is detailed at top right of each plot. Ovals represent 95% confidence inertia ellipses

Bibliography

Abaad, M. *et al.* (2016) 'Phenotypic plasticity in wild marine fishes associated with fish-cage aquaculture', *Hydrobiologia*, 765(1), pp. 343–358. doi: 10.1007/s10750-015-2428-5.

Agrawal, A. A., Laforsch, C. and Tollrian, R. (1999) 'Transgenerational induction of defences in animals and plants', *Nature*, 401(6748), pp. 60–63. doi: 10.1038/43425.

Albokhadaim, I. *et al.* (2007) 'Larval programming of post-hatch muscle growth and activity in Atlantic salmon (*Salmo salar*)', *Journal of Experimental Biology*, 210(10), pp. 1735–1741. doi: 10.1242/jeb.003194.

Allen, G. (1985) 'Snappers of the world: an annotated and illustrated catalogue of the Lutjanid species known to date.', *FAO Fisheries Synopsis*, 12, pp. 1–208.

Allendorf, F. W., Hohenlohe, P. A. and Luikart, G. (2010) 'Genomics and the future of conservation genetics.', *Nature reviews. Genetics*, 11, pp. 697–709. doi: 10.1038/nrg2844.

Allison, G. W. *et al.* (2003) 'ENSURING PERSISTENCE OF MARINE RESERVES: CATASTROPHES REQUIRE ADOPTING AN INSURANCE FACTOR', *Ecological Applications*. John Wiley & Sons, Ltd, 13(sp1), pp. 8–24. doi: 10.1890/1051-0761(2003)013[0008:EPOMRC]2.0.CO;2.

Almany, G. R. *et al.* (2007) 'Local replenishment of coral reef fish populations in a marine reserve.', *Science (New York, N.Y.)*. United States, 316(5825), pp. 742–744. doi: 10.1126/science.1140597.

Anderson, W. . and Allen, G. R. 2001 (2001) *Lutjanidae. Jobfishes*. Rome, Italy.

Andrew, S. and Marco, D. G. (2012) 'Linking behavioural syndromes and cognition: a behavioural ecology perspective', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 367(1603), pp. 2762–2772. doi: 10.1098/rstb.2012.0216.

Arnett, H. A. and Kinnison, M. T. (2016) 'Predator-induced phenotypic plasticity of shape and behavior: parallel and unique patterns across sexes and species', *Current Zoology*, 63(4), pp. 369–378. doi: 10.1093/cz/zow072.

B. Johnson, J. and Basolo, A. (2003) *Predator exposure alters female mate choice in the green swordtail*, *Behavioral Ecology*. doi: 10.1093/beheco/arg046.

Baird, N. A. *et al.* (2008) 'Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers', *PLoS ONE*. Edited by J. C. Fay. San Francisco, USA: Public Library of Science, 3(10), p. e3376. doi:

10.1371/journal.pone.0003376.

Barley, Shanta C, Meekan, M. G. and Meeuwig, J. J. (2017) 'Diet and condition of mesopredators on coral reefs in relation to shark abundance', *PLOS ONE*. Public Library of Science, 12(4), p. e0165113. Available at: <https://doi.org/10.1371/journal.pone.0165113>.

Barley, Shanta C., Meekan, M. G. and Meeuwig, J. J. (2017) 'Species diversity, abundance, biomass, size and trophic structure of fish on coral reefs in relation to shark abundance', *Marine Ecology Progress Series*, 565(February), pp. 163–179. doi: 10.3354/meps11981.

Baum, J. K. and Myers, R. A. (2004) 'Shifting baselines and the decline of pelagic sharks in the Gulf of Mexico', *Ecology Letters*. Blackwell Science Ltd, 7(2), pp. 135–145. doi: 10.1111/j.1461-0248.2003.00564.x.

Bekkevold, D. *et al.* (2005) 'ENVIRONMENTAL CORRELATES OF POPULATION DIFFERENTIATION IN ATLANTIC HERRING', *Evolution*, 59(12), pp. 2656–2668. doi: 10.1111/j.0014-3820.2005.tb00977.x.

Bellwood, D. R. *et al.* (2004) 'Confronting the coral reef crisis.', *Nature*, 429(6994), pp. 827–833. doi: 10.1038/nature02691.

Benestan, L. *et al.* (2015) 'RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*)', *Molecular Ecology*, 24(13), pp. 3299–3315. doi: 10.1111/mec.13245.

Bethea, D. M., Buckel, J. A. and Carlson, J. K. (2004) 'Foraging ecology of the early life stages of four sympatric shark species', *Marine Ecology Progress Series*, 268(October), pp. 245–264. doi: 10.3354/meps268245.

Böhm, M. *et al.* (2013) 'The conservation status of the world's reptiles', *Biological Conservation*, 157, pp. 372–385. doi: 10.1016/j.biocon.2012.07.015.

Bonjean, F. and Lagerloef, G. S. E. (2002) 'Diagnostic Model and Analysis of the Surface Currents in the Tropical Pacific Ocean', *Journal of Physical Oceanography*, 32(10), pp. 2938–2954. doi: 10.1175/1520-0485(2002)032<2938:DMAAOT>2.0.CO;2.

Booth, G., Niccolucci, M. and Schuster, E. (1994) *Identifying Proxy sets in Multiple Linear Regression: An Aid to Better Coefficient Interpretation*, Forest Service, Rocky Mountain Dept of Agriculture.

Borcherding, J. and Magnhagen, C. (2008) 'Food abundance affects both morphology and behaviour of juvenile perch', *Ecology of Freshwater Fish*. John Wiley & Sons, Ltd, 17(2), pp. 207–218. doi: <https://doi.org/10.1111/j.1600-0633.2007.00272.x>.

Bretman, A., Gage, M. J. G. and Chapman, T. (2011) 'Quick-change artists: male plastic behavioural responses to rivals', *Trends in Ecology & Evolution*, 26(9), pp. 467–473. doi: <https://doi.org/10.1016/j.tree.2011.05.002>.

Carlton, R., Dempsey, A., Lubarsky, K., Faisal, M., and Purkis, S. (2021) *Global Reef Expedition: Chagos Archipelago. Final Report(2021) Khaled bin Sultan Living Oceans Foundation, Annapolis, MD. Vol 13.*

Catano, L. B. *et al.* (2016) 'Reefscapes of fear: predation risk and reef heterogeneity interact to shape herbivore foraging behaviour.', *The Journal of animal ecology*. England, 85(1), pp. 146–156. doi: 10.1111/1365-2656.12440.

Catchen, J. *et al.* (2013) 'Stacks: an analysis tool set for population genomics', *Molecular Ecology*, 22(11), pp. 3124–3140. doi: 10.1111/mec.12354.

Catchen, J. M. *et al.* (2011) 'Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences', *G3: Genes/Genomes/Genetics*, 1(3), pp. 171–182. doi: 10.1534/g3.111.000240.

Chalfoun, A. D. and Martin, T. E. (2010) 'Facultative nest patch shifts in response to nest predation risk in the Brewer's sparrow: a "win-stay, lose-switch" strategy?', *Oecologia*. Germany, 163(4), pp. 885–892. doi: 10.1007/s00442-010-1679-0.

Chapman, L. J. (2015) 'Low-Oxygen Lifestyles BT - Extremophile Fishes: Ecology, Evolution, and Physiology of Teleosts in Extreme Environments', in Riesch, R., Tobler, M., and Plath, M. (eds). Cham: Springer International Publishing, pp. 9–33. doi: 10.1007/978-3-319-13362-1_2.

Chevalier, C. *et al.* (2017) 'The atypical hydrodynamics of the Mayotte Lagoon (Indian Ocean): Effects on water age and potential impact on plankton productivity', *Estuarine, Coastal and Shelf Science*, 196. doi: 10.1016/j.ecss.2017.06.027.

CHRISTIE, M. R. *et al.* (2010) 'Self-recruitment and sweepstakes reproduction amid extensive gene flow in a coral-reef fish', *Molecular Ecology*. John Wiley & Sons, Ltd (10.1111), 19(5), pp. 1042–1057. doi: 10.1111/j.1365-294X.2010.04524.x.

Christou, M. *et al.* (2018) 'Segmentation pattern of zebrafish caudal fin is affected by developmental temperature and defined by multiple fusions between segments', *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*. John Wiley & Sons, Ltd, 330(6–7), pp. 330–340. doi: <https://doi.org/10.1002/jez.b.22825>.

Cimmaruta, R. *et al.* (2008) 'Genetic structure and temporal stability in the horse mackerel (*Trachurus trachurus*)', *Fisheries Research*, 89(2), pp. 114–121. doi: <https://doi.org/10.1016/j.fishres.2007.09.030>.

Claro, R. *et al.* (2003) 'Gulf and Caribbean Research Spawning Aggregation Sites of Snapper and Grouper Species (Lutjanidae and Serranidae) on the Insular Shelf of Cuba SPAWNING AGGREGATION SITES OF SNAPPER AND GROUPER SPECIES (LUTJANIDAE AND SERRANIDAE) ON THE INSULAR SHELF O', 14(2), pp. 91–106.

Collen, B. *et al.* (2014) 'Global patterns of freshwater species diversity, threat and endemism', *Global Ecology and Biogeography*, 23(1), pp. 40–51. doi: 10.1111/geb.12096.

Collin, H. and Fumagalli, L. (2011) 'Evidence for morphological and adaptive genetic divergence between lake and stream habitats in European minnows (*Phoxinus phoxinus*, Cyprinidae).', *Molecular ecology*. England, 20(21), pp. 4490–4502. doi: 10.1111/j.1365-294X.2011.05284.x.

Cook, R. M., Sinclair, A. and Stefánsson, G. (1997) 'Potential collapse of North Sea cod stocks', *Nature*, 385(6616), pp. 521–522. doi: 10.1038/385521a0.

Couper, A. (1987) *The Times Atlas of The Oceans*. Van Nostrand Reinhold.

Cowen, R. (2002) 'Larval Dispersal and Retention and Consequences for Population Connectivity', in *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem*, pp. 149–I. doi: 10.1016/B978-012615185-5/50010-4.

Cowen, R. K. and Sponaugle, S. (2009) 'Larval Dispersal and Marine Population Connectivity', *Annual Review of Marine Science*. Annual Reviews, 1(1), pp. 443–466. doi: 10.1146/annurev.marine.010908.163757.

Creel, S. and Christianson, D. (2008) 'Relationships between direct predation and risk effects.', *Trends in ecology & evolution*. England, 23(4), pp. 194–201. doi: 10.1016/j.tree.2007.12.004.

Crozier, L. G. and Hutchings, J. A. (2014) 'Plastic and evolutionary responses to climate change in fish', *Evolutionary Applications*, 7(1), pp. 68–87. doi: doi:10.1111/eva.12135.

Currens, K. P. *et al.* (1989) 'Effects of different feeding regimes on the morphometrics of chinook salmon (*Oncorhynchus tshawytscha*) and rainbow trout (*O. mykiss*)', *Copeia*, 1989, pp. 689–695.

Davey, J. W. *et al.* (2011) 'Genome-wide genetic marker discovery and genotyping using next-generation sequencing.', *Nature reviews. Genetics*. England, 12(7), pp. 499–510. doi: 10.1038/nrg3012.

Davis, A. M. *et al.* (2012) 'Marine–freshwater transitions are associated with the evolution of dietary diversification in terapontid grunters (Teleostei: Terapontidae)', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, 25(6), pp. 1163–1179. doi: <https://doi.org/10.1111/j.1420-9101.2012.02504.x>.

Davis, A. M., Pusey, B. J. and Pearson, R. G. (2012) 'Contrasting intraspecific dietary shifts in two terapontid assemblages from Australia's wet-dry tropics', *Ecology of Freshwater Fish*. John Wiley & Sons, Ltd, 21(1), pp. 42–56. doi: <https://doi.org/10.1111/j.1600-0633.2011.00521.x>.

Denver, R. J., Mirhadi, N. and Phillips, M. (1998) 'ADAPTIVE PLASTICITY IN AMPHIBIAN METAMORPHOSIS: RESPONSE OF SCAPHIOPUS HAMMONDI TADPOLES TO HABITAT DESICCATION', *Ecology*. John Wiley & Sons, Ltd, 79(6), pp. 1859–1872. doi: 10.1890/0012-9658(1998)079[1859:APIAMR]2.0.CO;2.

Dewitt, T. J., Robinson, B. W. and Wilson, D. S. (2000) 'Functional diversity among predators of a freshwater snail imposes an adaptive trade-off for shell morphology DeWitt et al. 2002 *Evol. Ecol. Res.*.pdf', pp. 129–148.

Dirzo, R. *et al.* (2014) 'Defaunation in the Anthropocene Rodolfo Dirzo', 401(6195). doi: 10.1126/science.1251817.

Dixon, D. L. *et al.* (2011) 'Terrestrial chemical cues help coral reef fish larvae locate settlement habitat surrounding islands', *Ecology and Evolution*. John Wiley & Sons, Ltd, 1(4), pp. 586–595. doi: <https://doi.org/10.1002/ece3.53>.

Domenici, P. *et al.* (2008) 'Predator-induced morphology enhances escape locomotion in crucian carp', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 275(1631), pp. 195–201. doi: 10.1098/rspb.2007.1088.

Drass, D. M. *et al.* (2000) 'Larval development of red snapper, *Lutjanus campechanus*, and comparisons with co-occurring snapper species', *Fishery Bulletin*, 98(3), pp. 507–527.

Dulvy, N. K. *et al.* (2004) 'Size structural change in lightly exploited coral reef fish communities: evidence for weak indirect effects', *Canadian Journal of Fisheries and Aquatic Sciences*. NRC Research Press, 61(3), pp. 466–475. doi: 10.1139/f03-169.

Dunne, R. P. *et al.* (2014) *The Creation of the Chagos Marine Protected Area. A Fisheries Perspective, Advances in Marine Biology*. doi: 10.1016/B978-0-12-800214-8.00003-7.

Eklöv, P. and Svanbäck, R. (2006) 'Predation Risk Influences Adaptive Morphological Variation in Fish Populations.', *The American Naturalist*. The University of Chicago Press, 167(3), pp. 440–452. doi: 10.1086/499544.

Endler, J. A. (1973) 'Gene Flow and Population Differentiation', *Science*, 179(4070), pp. 243 LP – 250. doi: 10.1126/science.179.4070.243.

Ernande, B., Dieckmann, U. and Heino, M. (2004) *Adaptive changes in harvested populations: Plasticity and evolution of age and size at maturation, Proceedings. Biological sciences / The Royal*

Society. doi: 10.1098/rspb.2003.2519.

Eschmeyer, W., Fricke, R. and Van der Laan, R. (2016) *Catalog of Fishes: Genera, Species, References*.

Etter, P. D. *et al.* (2011) 'SNP Discovery and Genotyping for Evolutionary Genetics Using RAD Sequencing', *Methods in molecular biology (Clifton, N.J.)*, 772, pp. 157–178. doi: 10.1007/978-1-61779-228-1_9.

Favaloro, E. and Mazzola, A. (2003) 'Shape change during the growth of sharpsnout seabream reared under different conditions in a fish farm of the southern Tyrrhenian Sea', *Aquacultural Engineering*, 29(1), pp. 57–63. doi: [https://doi.org/10.1016/S0144-8609\(03\)00050-5](https://doi.org/10.1016/S0144-8609(03)00050-5).

Ferrario, F. *et al.* (2014) 'The effectiveness of coral reefs for coastal hazard risk reduction and adaptation', *Nature Communications*. Nature Publishing Group, 5(May), pp. 1–9. doi: 10.1038/ncomms4794.

Ferretti, F. *et al.* (2010) 'Patterns and ecosystem consequences of shark declines in the ocean', *Ecology Letters*, 13(8), pp. 1055–1071. doi: 10.1111/j.1461-0248.2010.01489.x.

Figuerola, D. E. (2008) '(Rajiformes : Rajidae : Dipturus) from Argentinean waters', *Zootaxa*, 46, p. 5326.

Fleminger, A. (1986) 'The Pleistocene equatorial barrier between the Indian and Pacific Oceans and a likely cause for Wallace's line', *UNESCO Tech.Pap.Mar.Sci*, 49, pp. 84–97.

Folkedal, O. *et al.* (2012) 'Duration of effects of acute environmental changes on food anticipatory behaviour, feed intake, oxygen consumption, and cortisol release in Atlantic salmon parr', *Physiology & Behavior*, 105(2), pp. 283–291. doi: <https://doi.org/10.1016/j.physbeh.2011.07.015>.

Fortaleza, M. and Nanola, C. (2017) 'Age Determination and Body Length Relationship of Two-Spot Red Snapper (*Lutjanus bohar*)', *Banwa B*, 12, pp. 1–12.

Froese R and D., P. (2017) *No Title, Fishbase*. Available at: www.fishbase.org (Accessed: 8 July 2017).

Gaffney, O. and Steffen, W. (2017) 'The Anthropocene equation', *Anthropocene Review*, 4(1), pp. 53–61. doi: 10.1177/2053019616688022.

Gaither, M. *et al.* (2009) 'Genetic evaluation of marine biogeographical barriers: Perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*)', *Journal of Biogeography - J BIOGEOGR*, 37, pp. 133–147. doi: 10.1111/j.1365-2699.2009.02188.x.

Garduno-Paz, M., Couderc, S. and Adams, C. (2010) 'Habitat complexity modulates phenotype expression through developmental plasticity in the threespine stickleback', *Biological Journal of the Linnean Society*, 100(2), pp. 407–413. doi: 10.1111/j.1095-8312.2010.01423.x.

Gerlach, G. *et al.* (2007) 'Smelling home can prevent dispersal of reef fish larvae', *Proceedings of the National Academy of Sciences*, 104(3), pp. 858 LP – 863. doi: 10.1073/pnas.0606777104.

Gillibrand, C. J., Harris, A. R. and Mara, E. (2007) 'Inventory and Spatial Assemblage Study of Reef Fish in the Area of Andavadoaka , South-West Madagascar (Western Indian Ocean)'.

Gomes, G., Sampaio, I. and Schneider, H. (2012) 'Population structure of *Lutjanus purpureus* (Lutjanidae - Perciformes) on the Brazilian coast: Further existence evidence of a single species of red snapper in the western Atlantic', *Anais da Academia Brasileira de Ciencias*, 84(4), pp. 979–999. doi: 10.1590/S0001-37652012000400013.

Graham, M. H. (2003) 'CONFRONTING MULTICOLLINEARITY IN ECOLOGICAL MULTIPLE REGRESSION', *Ecology*. John Wiley & Sons, Ltd, 84(11), pp. 2809–2815. doi: <https://doi.org/10.1890/02-3114>.

Graham, N. A. J. *et al.* (2013) 'The Status of Coral Reef Fish Assemblages in the Chagos Archipelago, with Implications for Protected Area Management and Climate Change.', in Sheppard, C. (ed.) *Coral Reefs of The World*. Springer, Dordrecht. doi: https://doi.org/10.1007/978-94-007-5965-7_19.

Graham, N. A. J., Nash, K. L. and Kool, J. T. (2011) 'Coral reef recovery dynamics in a changing world', *Coral Reefs*, 30(2), pp. 283–294. doi: 10.1007/s00338-010-0717-z.

Graham, N. A. J., Spalding, M. D. and Sheppard, C. R. C. (2010) 'Reef shark declines in remote atolls highlight the need for multi-faceted conservation action', *Aquatic Conservation: Marine and Freshwater Ecosystems*. John Wiley & Sons, Ltd., 20(5), pp. 543–548. doi: 10.1002/aqc.1116.

Grandcourt, E. M., Abdessalaam, T. Z. A. and Francis, F. (2006) 'Age, growth, mortality and reproduction of the blackspot snapper, *Lutjanus fulvivflamma* (Forsskal, 1775), in the southern Arabian Gulf', *Fisheries Research*, 78(2–3), pp. 203–210. doi: 10.1016/j.fishres.2005.11.021.

Green, A. L. *et al.* (2015) 'Larval dispersal and movement patterns of coral reef fishes , and implications for marine reserve network design', 90, pp. 1215–1247. doi: 10.1111/brv.12155.

Green, R. H., Lowe, R. J. and Buckley, M. L. (2018) 'Hydrodynamics of a Tidally Forced Coral Reef Atoll', *Journal of Geophysical Research: Oceans*. John Wiley & Sons, Ltd, 123(10), pp. 7084–7101. doi: 10.1029/2018JC013946.

Gusen, A. (2010) 'Environmental variation and phenotypic plasticity in perch (*Perca fluviatilis* L .)', pp. 1–25.

Hammerschlag, N. *et al.* (2018) 'Predator declines and morphological changes in prey: Evidence from coral reefs depleted of sharks', *Marine Ecology Progress Series*, 586, pp. 127–139. doi: 10.3354/meps12426.

Hanski, I. (1998) 'Metapopulation dynamics', *Nature*, 396(6706), pp. 41–49. doi: 10.1038/23876.

Harding, G. C. *et al.* (2011) 'Genetic relationships among subpopulations of the American lobster (*Homarus americanus*) as revealed by random amplified polymorphic DNA', *Canadian Journal of Fisheries and Aquatic Sciences*, 54(8), pp. 1762–1771. doi: 10.1139/f97-082.

Harrison, H. B. *et al.* (2012) 'Report Larval Export from Marine Reserves and the Recruitment Benefit for Fish and Fisheries', *Current Biology*. Elsevier Ltd, 22(11), pp. 1023–1028. doi: 10.1016/j.cub.2012.04.008.

Harrison, R. G. (1989) 'Animal mitochondrial DNA as a genetic marker in population and evolutionary biology', *Trends in Ecology & Evolution*, 4(1), pp. 6–11. doi: [https://doi.org/10.1016/0169-5347\(89\)90006-2](https://doi.org/10.1016/0169-5347(89)90006-2).

Harvell, C. D. *et al.* (2002) 'Climate warming and disease risks for terrestrial and marine biota.', *Science (New York, N.Y.)*. United States, 296(5576), pp. 2158–2162. doi: 10.1126/science.1063699.

Harvell, C. D. *et al.* (2008) 'Change and Ocean Acidification', (December 2007), pp. 1737–1743.

Hastings, A. (2004) *Hastings, A. Transients: the key to long-term ecological understanding? Trends Ecol. Evol.* 19, 39-45, *Trends in ecology & evolution*. doi: 10.1016/j.tree.2003.09.007.

Hauser Lorenz and Carvalho Gary R (2008) 'Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts', *Fish and Fisheries*, 9(4), pp. 333–362. doi: 10.1111/j.1467-2979.2008.00299.x.

Head, C. E. I., Bayley, D. T. I., Rowlands, G., Roche, R. C., Tickler, D. M., Rogers, A. D., Koldewey, H., Turner, J. R., *et al.* (2019) 'Coral bleaching impacts from back-to-back 2015–2016 thermal anomalies in the remote central Indian Ocean', *Coral Reefs*, 38(4), pp. 605–618. doi: 10.1007/s00338-019-01821-9.

Head, C. E. I., Bayley, D. T. I., Rowlands, G., Roche, R. C., Tickler, D. M., Rogers, A. D., Koldewey, H. and Andradi-brown, J. R. T. D. A. (2019) 'Coral bleaching impacts from back-to-back 2015 – 2016 thermal anomalies in the remote central Indian Ocean', *Coral Reefs*. Springer Berlin Heidelberg, 38(4), pp. 605–618. doi: 10.1007/s00338-019-01821-9.

Hebert, P. D. N., Ratnasingham, S. and de Waard, J. R. (2003) 'Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species', *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(Suppl 1), p. S96 LP-S99. Available at:

http://rspb.royalsocietypublishing.org/content/270/Suppl_1/S96.abstract.

HEDGECOCK, D., BARBER, P. H. and EDMANDS, S. (2007) 'Genetic Approaches to Measuring Connectivity', *Oceanography*. Oceanography Society, 20(3), pp. 70–79. Available at: <http://www.jstor.org/stable/24860097>.

Hegrenes, S. (2001) 'Diet-induced phenotypic plasticity of feeding morphology in the orangespotted sunfish, *Lepomis humilis*', *Ecology of Freshwater Fish*. John Wiley & Sons, Ltd, 10(1), pp. 35–42. doi: <https://doi.org/10.1034/j.1600-0633.2001.100105.x>.

Heithaus, M. R. *et al.* (2008) 'A review of lethal and non-lethal effects of predators on adult marine turtles', *Journal of Experimental Marine Biology and Ecology*, 356(1–2), pp. 43–51. doi: 10.1016/j.jembe.2007.12.013.

Hemmer-Hansen, J. *et al.* (2014) 'Population genomics of marine fishes: next generation prospects and challenges Population Genomics of Marine Fishes: Next-Generation Prospects and Challenges', *Biological Bulletin*, 227(2), pp. 117–132. doi: 10.1086/BBLv227n2p117.

Hendry, A. P., Farrugia, T. J. and Kinnison, M. T. (2008) 'Human influences on rates of phenotypic change in wild animal populations.', *Molecular ecology*. England, 17(1), pp. 20–29. doi: 10.1111/j.1365-294X.2007.03428.x.

Heron, S. F. *et al.* (2016) 'Local-scale projections of coral reef futures and implications of the Paris Agreement', *Scientific Reports*. Nature Publishing Group, 6(1), pp. 1–8. doi: 10.1038/srep39666.

Herrón, P. and Zapata, F. (2002) 'Pelagic larval duration and geographic distribution of tropical eastern Pacific snappers (Pisces: Lutjanidae)', *Marine Ecology Progress Series*, 230, pp. 295–300. doi: 10.3354/meps230295.

Hess, J. E. *et al.* (2013) 'Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species', *Molecular Ecology*. John Wiley & Sons, Ltd (10.1111), 22(11), pp. 2898–2916. doi: 10.1111/mec.12150.

Hochachka, P., Peter, W. and Somero, G. (2002) *Biochemical adaptation: mechanisms and processes in physiological evolution*. Oxford University Press.

Hoffmann, A. A. and Sgrò, C. M. (2011) 'Climate change and evolutionary adaptation', *Nature*, 470(7335), pp. 479–485. doi: 10.1038/nature09670.

Hollenbeck, C. M. *et al.* (2015) 'Population structure of red snapper (*Lutjanus campechanus*) in U.S. waters of the western Atlantic Ocean and the northeastern Gulf of Mexico', *Fisheries Research*, 172, pp. 17–25. doi: <https://doi.org/10.1016/j.fishres.2015.06.020>.

Van Hooidonk, R. *et al.* (2014) 'Opposite latitudinal gradients in projected ocean acidification and bleaching impacts on coral reefs', *Global Change Biology*, 20(1), pp. 103–112. doi: 10.1111/gcb.12394.

Hutchinson, W. F. *et al.* (2003) 'Temporal analysis of archived samples indicates marked genetic changes in declining North Sea cod (*Gadus morhua*)', *Proceedings of the Royal Society B: Biological Sciences*, 270(1529), pp. 2125–2132. doi: 10.1098/rspb.2003.2493.

J, F. A. *et al.* (2007) 'Plasticity in animal personality traits: does prior experience alter the degree of boldness?', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 274(1608), pp. 333–339. doi: 10.1098/rspb.2006.3751.

Jackson, A. M. *et al.* (2014) 'Population Structure and Phylogeography in Nassau Grouper (*Epinephelus striatus*), a Mass-Aggregating Marine Fish', *PLoS ONE*. Edited by G. H. Yue. Public Library of Science, 9(5), p. e97508. doi: 10.1371/journal.pone.0097508.

Jackson, J. B. C. *et al.* (2001) 'Historical Collapse Overfishing of and the Recent Coastal Ecosystems', *Science*, 293(5530), pp. 629–638. doi: 10.1126/science.1059199.

Jacobson, P. *et al.* (2018) 'Size-dependent prey availability affects diet and performance of predatory fish at sea: a case study of Atlantic salmon', *Ecosphere*. John Wiley & Sons, Ltd, 9(1), p. e02081. doi: <https://doi.org/10.1002/ecs2.2081>.

Jahn, A. (2018) 'Reduced probability of ice-free summers for 1.5 °C compared to 2 °C warming', *Nature Climate Change*. Springer US, 8(5), pp. 409–413. doi: 10.1038/s41558-018-0127-8.

Jaxion-Harm, J. and Ladich, F. (2014) 'Effects of temperature change on cortisol release by common carp *Cyprinus carpio*', *Journal of Fish Biology*. John Wiley & Sons, Ltd, 84(4), pp. 1221–1227. doi: <https://doi.org/10.1111/jfb.12331>.

Jennings, S., Kaiser, M. J. and Reynolds, J. D. (2001) *Marine Fisheries Ecology*. Blackwell Publishing, Malden, MA.

Johnston, I. and Bennett, A. (2008) *Animals and temperature: phenotypic plasticity and evolutionary adaptation*. Cambridge University Press.

Jombart, T. and Ahmed, I. (2011) 'adegenet 1.3-1: new tools for the analysis of genome-wide SNP data', *Bioinformatics*, 27(21), pp. 3070–3071. doi: 10.1093/bioinformatics/btr521.

Jonsson, B. and Jonsson, N. (2014) 'Early environment influences later performance in fishes', *Journal of Fish Biology*. John Wiley & Sons, Ltd, 85(2), pp. 151–188. doi: <https://doi.org/10.1111/jfb.12432>.

Jorde, P. E. *et al.* (2003) 'Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod', *Molecular Ecology*, 12, pp. 385–394.

Kausar, R. and Salim, M. (2006) 'Effect of Water Temperature on the Growth Performance and Feed Conversion Ratio of Labeo Rohita', *Pakistan Veterinary Journal*, 26(3), pp. 105–108.

Kelly, S. A., Panhuis, T. M. and Stoeckl, A. M. (2012) 'Phenotypic plasticity: Molecular mechanisms and adaptive significance', *Comprehensive Physiology*, 2(2), pp. 1417–1439. doi: 10.1002/cphy.c110008.

Kenchington, E. *et al.* (2009) *Pleistocene glaciation events shape genetic structure across the range of the American lobster, Homarus americanus*, *Molecular ecology*. doi: 10.1111/j.1365-294X.2009.04118.x.

King, A. J. *et al.* (2013) 'Sex-Differences and Temporal Consistency in Stickleback Fish Boldness', *PLOS ONE*. Public Library of Science, 8(12), p. e81116. Available at: <https://doi.org/10.1371/journal.pone.0081116>.

Kingsford, M. J. *et al.* (2002) 'Sensory environments, larval abilities and local self-recruitment', *Bulletin of Marine Science*, 70(1 SUPPL.), pp. 309–340.

Knutsen, H. *et al.* (2011) 'Are low but statistically significant levels of genetic differentiation in marine fishes "biologically meaningful"? A case study of coastal Atlantic cod', *Molecular Ecology*, 20(4), pp. 768–783. doi: 10.1111/j.1365-294X.2010.04979.x.

Kochzius, M. *et al.* (2010) 'Identifying fishes through DNA barcodes and microarrays', *PLoS ONE*, 5(9), pp. 1–15. doi: 10.1371/journal.pone.0012620.

Kröger, R. H. H. and Fernald, R. D. (1994) 'Regulation of eye growth in the African cichlid fish *Haplochromis burtoni*', *Vision Research*, 34(14), pp. 1807–1814. doi: [https://doi.org/10.1016/0042-6989\(94\)90305-0](https://doi.org/10.1016/0042-6989(94)90305-0).

L. Tracey, M. *et al.* (2011) *Biochemical Genetics of Lobsters: Genetic Variation and the Structure of American Lobster (Homarus americanus) Populations*, *Journal of the Fisheries Research Board of Canada*. doi: 10.1139/f75-247.

Lacasse, J. and Aubin-Horth, N. (2015) 'A test of the coupling of predator defense morphology and behavior variation in two threespine stickleback populations', *Current Zoology*, 58(1), pp. 53–65. doi: 10.1093/czoolo/58.1.53.

LANDE, R. (2009) 'Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, 22(7), pp. 1435–1446. doi: <https://doi.org/10.1111/j.1420-9101.2009.01754.x>.

- Langerhans, R. B. *et al.* (2004) 'Predator-Driven Phenotypic Diversification in *Gambusia affinis* PREDATOR-DRIVEN PHENOTYPIC DIVERSIFICATION IN GAMBUSIA AFFINIS', 58(10), pp. 2305–2318.
- Larson, W. A. *et al.* (2014) 'Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*)', *Evolutionary Applications*, 7(3), pp. 355–369. doi: 10.1111/eva.12128.
- Law, R. (2000) 'Fishing, selection, and phenotypic evolution', *ICES Journal of Marine Science*, 57(3), pp. 659–668. doi: 10.1006/jmsc.2000.0731.
- Lee, W.-S., Monaghan, P. and Metcalfe, N. B. (2013) 'Experimental demonstration of the growth rate–lifespan trade-off', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 280(1752), p. 20122370. doi: 10.1098/rspb.2012.2370.
- Leis, J. (2010) 'Ontogeny of behaviour in larvae of marine demersal fishes', *Ichthyological Research*, 57, pp. 325–342. doi: 10.1007/s10228-010-0177-z.
- Lever, J., Krzywinski, M. and Altman, N. (2017) 'Points of Significance: Principal component analysis', *Nature Methods*. Nature Publishing Group, 14(7), pp. 641–642. doi: 10.1038/nmeth.4346.
- M, de R. A., S, B. D. and Lennart, P. (2006) 'Evolutionary regime shifts in age and size at maturation of exploited fish stocks', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society, 273(1596), pp. 1873–1880. doi: 10.1098/rspb.2006.3518.
- Ma, Y., Kjesbu, O. S. and Jørgensen, T. (1998) 'Effects of ration on the maturation and fecundity in captive Atlantic herring (*Clupea harengus*)', *Canadian Journal of Fisheries and Aquatic Sciences*. NRC Research Press, 55(4), pp. 900–908. doi: 10.1139/f97-305.
- Mace, G. and Purvis, A. (2008) 'Evolutionary biology and practical conservation: bridging a widening gap', *Molecular Ecology*. John Wiley & Sons, Ltd, 17(1), pp. 9–19. doi: <https://doi.org/10.1111/j.1365-294X.2007.03455.x>.
- Madin, E. M. P., Madin, J. S. and Booth, D. J. (2011) 'Landscape of fear visible from space', *Scientific Reports*, 1(1), p. 14. doi: 10.1038/srep00014.
- Madison, B. N. *et al.* (2015) 'Chronic cortisol and the regulation of food intake and the endocrine growth axis in rainbow trout', *Journal of Endocrinology*, 226(2), pp. 103–119. doi: 10.1530/JOE-15-0186.
- Marriott, R. J., Mapstone, B. D. and Begg, G. A. (2007) 'Age-specific demographic parameters, and their implications for management of the red bass, *Lutjanus bohar* (Forsskal 1775): A large, long-lived reef fish', *Fisheries Research*, 83(2), pp. 204–215. doi: <https://doi.org/10.1016/j.fishres.2006.09.016>.

Marteinsdottir, G. and Steinarsson, A. (1998) 'Maternal influence on the size and viability of Iceland cod *Gadus morhua* eggs and larvae', *Journal of Fish Biology*. Blackwell Publishing Ltd, 52(6), pp. 1241–1258. doi: 10.1111/j.1095-8649.1998.tb00969.x.

Mateus, A. P. *et al.* (2017) 'Thermal imprinting modifies adult stress and innate immune responsiveness in the teleost sea bream.', *The Journal of endocrinology*. England, 233(3), pp. 381–394. doi: 10.1530/JOE-16-0610.

Matzke, N. *et al.* (2011) 'Has the Earth's sixth mass extinction already arrived?', *Nature*. Nature Publishing Group, 471(7336), pp. 51–57. doi: 10.1038/nature09678.

Mihalitsis, M. and Bellwood, D. R. (2019) 'Morphological and functional diversity of piscivorous fishes on coral reefs', *Coral Reefs*, 38(5), pp. 945–954. doi: 10.1007/s00338-019-01820-w.

Mitchem, L. D. *et al.* (2018) 'The pervasive effects of lighting environments on sensory drive in bluefin killifish: an investigation into male/male competition, female choice, and predation', *Current zoology*. 2018/05/31. Oxford University Press, 64(4), pp. 499–512. doi: 10.1093/cz/zoy038.

Moore, B. (2019) 'Age-based life history of humpback red snapper, *Lutjanus gibbus*, in New Caledonia', *Journal of Fish Biology*. doi: 10.1111/jfb.14142.

Moran, P., Bradbury, R. and Reichelt, R. (1988) 'Distribution of recent outbreaks of the crown of thorns starfish (*Acanthaster planci*) along the Great Barrier Reef', *Coral Reefs*, 7, pp. 125–137.

Moritz, C., Dowling, T. E. and Brown, W. M. (1987) 'Evolution of animal mitochondrial DNA: relevance for population biology and systematics.', *Annual review of ecology and systematics*. Vol. 18, 92, pp. 269–292. doi: 10.1146/annurev.es.18.110187.001413.

MRAG (2014) 'Strengthening enforcement in the British Indian Ocean Territory', *Tech. Rep.*

Muths, D. *et al.* (2012) 'Genetic connectivity of the reef fish *Lutjanus kasmira* at the scale of the western Indian Ocean', 69(May), pp. 842–853.

Myers, R. A., Hutchings, J. A. and Barrowman, N. J. (1996) 'Hypotheses for the decline of cod in the North Atlantic', *Marine Ecology Progress Series*, 138(1–3), pp. 293–308. doi: 10.3354/meps138293.

Nagelkerken, I. *et al.* (2000) 'Day-night shifts of fishes between shallow-water biotopes of a Caribbean bay, with emphasis on the nocturnal feeding of Haemulidae and Lutjanidae', *Marine Ecology Progress Series*, 194, pp. 55–64. doi: 10.3354/meps194055.

Nagelkerken, I *et al.* (2000) 'Day-night shifts of fishes between shallow-water biotopes of a Caribbean bay, with emphasis on the nocturnal feeding of Haemulidae and Lutjanidae ', *Marine*

Ecology Progress Series, 194, pp. 55–64. Available at: <https://www.int-res.com/abstracts/meps/v194/p55-64/>.

Naish, K. A. and Hard, J. J. (2008) 'Bridging the gap between the genotype and the phenotype: linking genetic variation, selection and adaptation in fishes', *Fish and Fisheries*. John Wiley & Sons, Ltd (10.1111), 9(4), pp. 396–422. doi: 10.1111/j.1467-2979.2008.00302.x.

Nanami, A. *et al.* (2010) 'Age, growth and reproduction of the humpback red snapper *Lutjanus gibbus* off Ishigaki Island, Okinawa', *Ichthyological Research*, 57(3), pp. 240–244. doi: 10.1007/s10228-010-0160-8.

Nanninga, G. B. and Spaet, J. L. Y. (2017) 'Spawning aggregations of the Humpback red snapper, *Lutjanus gibbus*, in the Tuamotus, French Polynesia', *Marine Biodiversity*, 47(2), pp. 375–376. doi: 10.1007/s12526-016-0497-y.

Narum, S. R. *et al.* (2013) 'Genotyping-by-sequencing in ecological and conservation genomics', *Molecular Ecology*. John Wiley & Sons, Ltd (10.1111), 22(11), pp. 2841–2847. doi: 10.1111/mec.12350.

Nei, M. (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Nielsen, E. E. *et al.* (2004) 'Genetic population structure of turbot (*Scophthalmus maximus* L.) supports the presence of multiple hybrid zones for marine fishes in the transition zone between the Baltic Sea and the North Sea', *Molecular Ecology*, 13(3), pp. 585–595. doi: 10.1046/j.1365-294X.2004.02097.x.

Nielsen, E. E. *et al.* (2009) 'Population genomics of marine fishes: identifying adaptive variation in space and time', *Molecular Ecology*. John Wiley & Sons, Ltd, 18(15), pp. 3128–3150. doi: 10.1111/j.1365-294X.2009.04272.x.

Ninio, R. *et al.* (2000) 'Temporal patterns in coral assemblages on the Great Barrier Reef from local to large spatial scales', *Marine Ecology Progress Series*, 194, pp. 65–74. doi: 10.3354/meps194065.

Nobre, C. A. *et al.* (2016) 'Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm', *Proceedings of the National Academy of Sciences*, 113(39), pp. 10759–10768. doi: 10.1073/pnas.1605516113.

Obura, D. (2012) 'The Diversity and Biogeography of Western Indian Ocean Reef-Building Corals', *PloS one*, 7, p. e45013. doi: 10.1371/journal.pone.0045013.

Okiyama, M. (1988) *An atlas of the early stage fishes in Japan*. Tokai University Press.

Olsen, E. *et al.* (2004) *Life-history variation among local populations of Atlantic cod from the Norwegian Skagerrak coast*, *Journal of Fish Biology - J FISH BIOL.* doi: 10.1111/j.0022-1112.2004.00402.x.

Olsen, E. M. *et al.* (2004) 'Maturation trends indicative of rapid evolution preceded the collapse of northern cod', *Nature*, 428(6986), pp. 932–935. doi: 10.1038/nature02430.

Olsen, E. M. *et al.* (2008) 'Small-scale biocomplexity in coastal Atlantic cod supporting a Darwinian perspective on fisheries management', *Evolutionary Applications*, 1(3), pp. 524–533. doi: 10.1111/j.1752-4571.2008.00024.x.

Olsson, J. and Eklöv, P. (2005) 'Habitat structure, feeding mode and morphological reversibility: Factors influencing phenotypic plasticity in perch', *Evolutionary Ecology Research*, 7(8), pp. 1109–1123.

Palkovacs, E. P., Wasserman, B. A. and Kinnison, M. T. (2011) 'Eco-evolutionary trophic dynamics: Loss of top predators drives trophic evolution and ecology of prey', *PLoS ONE*, 6(4). doi: 10.1371/journal.pone.0018879.

Palumbi, S. R. (1994) 'GENETIC DIVERGENCE, REPRODUCTIVE ISOLATION, AND MARINE SPECIATION', *Annual Review of Ecology and Systematics*. Annual Reviews, 25(1), pp. 547–572. doi: 10.1146/annurev.es.25.110194.002555.

Pampoulie, C. *et al.* (2011) *The genetic structure of Atlantic cod (Gadus morhua) around Iceland: insight from microsatellites, the Pan I locus, and tagging experiments*, *Canadian Journal of Fisheries and Aquatic Sciences*. doi: 10.1139/f06-150.

Pandolfi, J. M. *et al.* (2011) 'Projecting Coral Reef Futures Under Global Warming and Ocean Acidification', *Science*, 333(6041), pp. 418–422. doi: 10.1126/science.1204794.

Papastamatiou, Y. P. *et al.* (2006) 'Distribution and diet of four species of carcharhinid shark in the Hawaiian Islands: Evidence for resource partitioning and competitive exclusion', *Marine Ecology Progress Series*, 320(August), pp. 239–251. doi: 10.3354/meps320239.

Papastamatiou, Y. P. *et al.* (2010) 'Long-term movement patterns and trophic ecology of blacktip reef sharks (*Carcharhinus melanopterus*) at Palmyra Atoll', *Journal of Experimental Marine Biology and Ecology*, 386(1–2), pp. 94–102. doi: 10.1016/j.jembe.2010.02.009.

Paris, C., Cherubin, L. and Cowen, R. (2007) 'Surfing, spinning, or diving from reef to reef: Effects on population connectivity', *Marine Ecology-progress Series - MAR ECOL-PROGR SER*, 347, pp. 285–300. doi: 10.3354/meps06985.

Park, J. Y. *et al.* (2016) 'Survival Rate and Hematological Responses with Temperature Changes of

Red Spotted Grouper, *Epinephelus akaara* in South Korea', *Development & reproduction*. The Korean Society of Developmental Biology, 20(2), pp. 103–112. doi: 10.12717/DR.2016.20.2.103.

Paul, M. *et al.* (2017) 'Relationships of body lengths with mouth opening and prey length of nemipterid fishes (Regan, 1913) in the Gulf of Thailand', *The Egyptian Journal of Aquatic Research*, 43(4), pp. 297–302. doi: <https://doi.org/10.1016/j.ejar.2017.11.001>.

Pauly, D. and Zeller, D. (2016) 'Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining', *Nature Communications*. Nature Publishing Group, 7, pp. 1–9. doi: 10.1038/ncomms10244.

Pigliucci, M. (2005) 'Evolution of phenotypic plasticity: where are we going now?', *Trends in Ecology & Evolution*, 20(9), pp. 481–486. doi: <https://doi.org/10.1016/j.tree.2005.06.001>.

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

Pugh, D. T. and Rayner, R. F. (1981) 'The tidal regimes of three indian ocean atolls and some ecological implications', *Estuarine, Coastal and Shelf Science*, 13(4), pp. 389–407. doi: 10.1016/S0302-3524(81)80036-9.

Puritz, J. B., Gold, J. R. and Portnoy, D. S. (2016) 'Fine-scale partitioning of genomic variation among recruits in an exploited fishery: causes and consequences', *Scientific Reports*, 6(1), p. 36095. doi: 10.1038/srep36095.

Puritz, J. B., Hollenbeck, C. M. and Gold, J. R. (2014) '*dDocent* : a RADseq, variant-calling pipeline designed for population genomics of non-model organisms', *PeerJ*, 2, p. e431. doi: 10.7717/peerj.431.

Ramler, D. *et al.* (2014) 'Nonlinear effects of temperature on body form and developmental canalization in the threespine stickleback', *Journal of Evolutionary Biology*. John Wiley & Sons, Ltd, 27(3), pp. 497–507. doi: <https://doi.org/10.1111/jeb.12311>.

Reaka-Kudla, M. L., Wilson, D. E. and Wilson, E. O. (1997) *Biodiversity II: Understanding and Protecting Our Biological Resources*. Joseph Henry Press.

Robert, J. (2007) *Report for MRAG. London: 2007. Preliminary characterization of bycatch by tuna longliners operating in the Chagos/BIOT area.*

Roberts, S. and Gavery, M. (2012) 'Is There a Relationship between DNA Methylation and Phenotypic Plasticity in Invertebrates? ', *Frontiers in Physiology* , p. 116. Available at: <https://www.frontiersin.org/article/10.3389/fphys.2011.00116>.

Rochette, N. C. and Catchen, J. M. (2017) 'Deriving genotypes from RAD-seq short-read data using Stacks', *Nature Protocols*. Nature Publishing Group, 12(12), pp. 2640–2659. doi: 10.1038/nprot.2017.123.

Rodríguez-Ezpeleta, N. *et al.* (2016) 'Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection', *Molecular Ecology Resources*. John Wiley & Sons, Ltd (10.1111), 16(4), pp. 991–1001. doi: 10.1111/1755-0998.12518.

Rodriguez-Ezpeleta, N., Álvarez, P. and Irigoien, X. (2017) 'Genetic Diversity and Connectivity in *Maurolicus muelleri* in the Bay of Biscay Inferred from Thousands of SNP Markers ', *Frontiers in Genetics* , p. 195. Available at: <https://www.frontiersin.org/article/10.3389/fgene.2017.00195>.

Roesti, M., Salzburger, W. and Berner, D. (2012) 'Uninformative polymorphisms bias genome scans for signatures of selection', *BMC Evolutionary Biology*. ???, 12(1), p. 1. doi: 10.1186/1471-2148-12-94.

Rousset, F. (1997) 'Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance', *Genetics*, 145(4), pp. 1219 LP – 1228. Available at: <http://www.genetics.org/content/145/4/1219.abstract>.

Ruppert, J. L. W. *et al.* (2013) 'Caught in the Middle: Combined Impacts of Shark Removal and Coral Loss on the Fish Communities of Coral Reefs', *PLoS ONE*, 8(9), pp. 1–9. doi: 10.1371/journal.pone.0074648.

Sakaue, J. *et al.* (2016) 'Temporal and spatial site sharing during spawning in snappers *symphorichthys spilurus* and *lutjanus bohar* (Pisces: Perciformes: Lutjanidae) in waters around peleliu island, Palau', *Zoological Studies*, 55(September 2016). doi: 10.6620/ZS.2016.55-44.

Sarà, M., Favalaro, E. and Mazzola, A. (1999) 'Comparative morphometrics of sharpsnout seabream (*Diplodus puntazzo* Cetti, 1777), reared in different conditions', *Aquacultural Engineering*, 19(3), pp. 195–209. doi: [https://doi.org/10.1016/S0144-8609\(98\)00052-1](https://doi.org/10.1016/S0144-8609(98)00052-1).

Sarvas, T. H. and Fevolden, S. E. (2005) 'Pantophysin (Pan I) locus divergence between inshore v. offshore and northern v. southern populations of Atlantic cod in the north-east Atlantic', *Journal of Fish Biology*. John Wiley & Sons, Ltd (10.1111), 67(2), pp. 444–469. doi: 10.1111/j.0022-1112.2005.00738.x.

Sheehan, E. V *et al.* (2019) 'The Effect of Deep Oceanic Flushing on Water Properties and Ecosystem Functioning Within Atolls in the British Indian Ocean Territory ', *Frontiers in Marine Science* , p. 512. Available at: <https://www.frontiersin.org/article/10.3389/fmars.2019.00512>.

Sheppard, C. R. C. (2003) 'Predicted recurrences of mass coral mortality in the Indian Ocean.', *Nature*, 425(6955), pp. 294–297. doi: 10.1038/nature01987.

Sheppard, C R C *et al.* (2012) 'Reefs and islands of the Chagos Archipelago, Indian Ocean: why it is the world's largest no-take marine protected area', *Aquatic conservation : marine and freshwater ecosystems*, 22(2), pp. 232–261. doi: 10.1002/aqc.1248.

Sheppard, C. R C *et al.* (2012) 'Reefs and islands of the Chagos Archipelago, Indian Ocean: Why it is the world's largest no-take marine protected area', *Aquatic Conservation: Marine and Freshwater Ecosystems*, 22(2), pp. 232–261. doi: 10.1002/aqc.1248.

da Silva, R. *et al.* (2016) 'Lack of Spatial Subdivision for the Snapper *Lutjanus purpureus* (Lutjanidae – Perciformes) from Southwest Atlantic Based on Multi-Locus Analyses', *PLOS ONE*. Public Library of Science, 11(8), p. e0161617. Available at: <https://doi.org/10.1371/journal.pone.0161617>.

Simpson, S. D., Meekan, M. G. and McCauley, R. D. (2004) 'Attraction of settlement-stage coral reef fishes to reef noise ', *Marine Ecology Progress Series*, 276, pp. 263–268. Available at: <https://www.int-res.com/abstracts/meps/v276/p263-268/>.

Smith, P. J. *et al.* (2008) 'Molecular analysis of Southern Ocean skates (*Bathyraja*) reveals a new species of Antarctic skate', *Journal of Fish Biology*, 73(5), pp. 1170–1182. doi: 10.1111/j.1095-8649.2008.01957.x.

Soeparno *et al.* (2012) 'Relationship between pelagic larval duration and abundance of tropical fishes on temperate coasts of Japan', *Journal of Fish Biology*, 80(2), pp. 346–357. doi: 10.1111/j.1095-8649.2011.03175.x.

Sollid, J. and Nilsson, G. E. (2006) 'Plasticity of respiratory structures — Adaptive remodeling of fish gills induced by ambient oxygen and temperature', *Respiratory Physiology & Neurobiology*, 154(1), pp. 241–251. doi: <https://doi.org/10.1016/j.resp.2006.02.006>.

Steffen, W. *et al.* (2015) 'The trajectory of the Anthropocene: The Great Acceleration', *The Anthropocene Review*. SAGE Publications, 2(1), pp. 81–98. doi: 10.1177/2053019614564785.

Stephenson, R. and Kenchington, E. (2000) 'Conserving fish stock structure is a critical aspect of preserving biodiversity.', *ICES C.M. 2000/Mini : 07*.

Sully, S. *et al.* (2019) 'A global analysis of coral bleaching over the past two decades', *Nature Communications*. Springer US, 10(1), pp. 1–5. doi: 10.1038/s41467-019-09238-2.

Trippel, E. A. (1995) 'Age at Maturity as a Stress Indicator in Fisheries: Biological processes related to reproduction in northwest Atlantic groundfish populations that have undergone declines', *BioScience*, 45(11), pp. 759–771. doi: 10.2307/1312628.

Underwood, J. N. *et al.* (2009) 'Ecologically relevant dispersal of corals on isolated reefs: implications for managing resilience', *Ecological Applications*, 19(1), pp. 18–29. doi: <https://doi.org/10.1890/07->

Underwood, J. N., Travers, M. J. and Gilmour, J. P. (2012) 'Subtle genetic structure reveals restricted connectivity among populations of a coral reef fish inhabiting remote atolls', *Ecology and Evolution*, 2(3), pp. 666–679. doi: 10.1002/ece3.80.

Vagner, M., Zambonino-Infante, J.-L. and Mazurais, D. (2019) 'Fish facing global change: are early stages the lifeline?', *Marine Environmental Research*, 147, pp. 159–178. doi: <https://doi.org/10.1016/j.marenvres.2019.04.005>.

Volkoff, H. and Rønnestad, I. (2020) 'Effects of temperature on feeding and digestive processes in fish', *Temperature*. Taylor & Francis, 7(4), pp. 307–320. doi: 10.1080/23328940.2020.1765950.

Walsh, M. R. *et al.* (2006) 'Maladaptive changes in multiple traits caused by fishing: impediments to population recovery', *Ecology Letters*. John Wiley & Sons, Ltd (10.1111), 9(2), pp. 142–148. doi: 10.1111/j.1461-0248.2005.00858.x.

Waples, R. S. (1998) 'Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species', *Journal of Heredity*, 89(5), pp. 438–450. doi: 10.1093/jhered/89.5.438.

Ward-Paige, C. A. *et al.* (2010) 'Large-scale absence of sharks on reefs in the greater-caribbean: A footprint of human pressures', *PLoS ONE*, 5(8). doi: 10.1371/journal.pone.0011968.

Ward, R. D. *et al.* (2005) 'DNA barcoding Australia's fish species.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360(September), pp. 1847–1857. doi: 10.1098/rstb.2005.1716.

Ward, R. D., Hanner, R. and Hebert, P. D. N. (2009) 'The campaign to DNA barcode all fishes, FISH-BOL.', *Journal of fish biology*. England, 74(2), pp. 329–356. doi: 10.1111/j.1095-8649.2008.02080.x.

Ward, R. D., Holmes, B. H. and Yearsley, G. K. (2008) 'DNA barcoding reveals a likely second species of Asian sea bass (barramundi) (*Lates calcarifer*)', *Journal of Fish Biology*, 72(2), pp. 458–463. doi: 10.1111/j.1095-8649.2007.01703.x.

Ward, R. D., Woodward, M. and Skibinski, D. O. F. (1994a) 'A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes', *Journal of Fish Biology*. John Wiley & Sons, Ltd (10.1111), 44(2), pp. 213–232. doi: 10.1111/j.1095-8649.1994.tb01200.x.

Ward, R. D., Woodward, M. and Skibinski, D. O. F. (1994b) 'A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes', *Journal of Fish Biology*. John Wiley & Sons, Ltd (10.1111), 44(2), pp. 213–232. doi: 10.1111/j.1095-8649.1994.tb01200.x.

- Webb, P. W. (1984) 'Body Form, Locomotion and Foraging in Aquatic Vertebrates¹', *American Zoologist*, 24(1), pp. 107–120. doi: 10.1093/icb/24.1.107.
- Wetherbee, B. M., Crow, G. L. and Lowe, C. G. (1997) 'Distribution, reproduction and diet of the gray reef shark *Carcharhinus amblyrhynchos* in Hawaii', *Marine Ecology Progress Series*, 151(1–3), pp. 181–189. doi: 10.3354/meps151181.
- White, C. *et al.* (2010) 'Ocean currents help explain population genetic structure', *Proceedings of the Royal Society B: Biological Sciences*, 277(1688), pp. 1685–1694. doi: 10.1098/rspb.2009.2214.
- Wilson, D. and McCormick, M. (1999) 'Microstructure of settlement-marks in the otoliths of tropical reef fishes', *Marine Biology*, 134, pp. 29–41. doi: 10.1007/s002270050522.
- Wirsing, A. J. *et al.* (2008) 'Seascapes of fear: evaluating sublethal predator effects experienced and generated by marine mammals', *Marine Mammal Science*. Blackwell Publishing Inc, 24(1), pp. 1–15. doi: 10.1111/j.1748-7692.2007.00167.x.
- Wirsing, A. J., Heithaus, M. R. and Dill, L. M. (2007) 'Living on the edge: dugongs prefer to forage in microhabitats that allow escape from rather than avoidance of predators', *Animal Behaviour*, 74(1), pp. 93–101. doi: 10.1016/j.anbehav.2006.11.016.
- Wisenden, B. D. *et al.* (2011) 'Maternal inheritance and exploratory-boldness behavioural syndrome in zebrafish', *Behaviour*. BRILL, 148(14), pp. 1443–1456. Available at: <http://www.jstor.org/stable/41444677>.
- Worm, B. *et al.* (2006) 'Impacts of biodiversity loss on ocean ecosystem services.', *Science (New York, N.Y.)*. United States, 314(5800), pp. 787–790. doi: 10.1126/science.1132294.
- Wright, P. A. and Turko, A. J. (2016) 'Amphibious fishes: evolution and phenotypic plasticity', *Journal of Experimental Biology*, 219(15), pp. 2245–2259. doi: 10.1242/jeb.126649.
- WWF/ZSL. *The Living Planet Index database*, <www.livingplanetindex.org> (2018).
- WWF/ZSL (2020) *LIVING PLANET REPORT 2020*.
- Young, T. P., Stanton, M. L. and Christian, C. E. (2003) 'Effects of natural and simulated herbivory on spine lengths of *Acacia drepanolobium* in Kenya', *Oikos*. John Wiley & Sons, Ltd, 101(1), pp. 171–179. doi: <https://doi.org/10.1034/j.1600-0706.2003.12067.x>.
- Zhang, J., Cai, Z. and Huang, L. (2006) 'Population genetic structure of crimson snapper *Lutjanus erythropterus* in East Asia, revealed by analysis of the mitochondrial control region', *ICES Journal of Marine Science*, 63(4), pp. 693–704. doi: 10.1016/j.icesjms.2006.01.004.

