Victoria University of Wellington

# The phylogeographic structure and genetic differentiation of *Polyprion oxygeneios* and *Polyprion americanus* in the Southern Hemisphere based on mitochondrial DNA control region sequences and microsatellite DNA genotyping

*Author: Natalie Stewart*

*Supervisors: Peter Ritchie, Corey Wakefield*

*A thesis submitted in fulfilment of the requirements for the degree of Master of Science*

*in the School of Biological Sciences*

November 2020

# *Abstract*

The genus *Polyprion*, commonly known as Wreckfish, is comprised of two species: *Polyprion oxygeneios*, limited to continental, insular, and seamount slopes in the Southern Hemisphere; and *P. americanus*, widely distributed in both the Northern and Southern Hemisphere. *Polyprion* support recreational and commercial fisheries off the Juan Fernandez Islands, Chile, Australia, and New Zealand, and although the fishery is relatively small in tonnage, its market value is high per kilo when compared to other species. *Polyprion* are long-lived, slow growing, and late to mature. These are life history characteristics that make a species vulnerable to over-fishing. Understanding the stock structure of *Polyprion* is important for fisheries management plans to be properly aligned to reproductive units and natural population boundaries. There has been speculation that more than two species of *Polyprion* exist in the Southern Hemisphere, which is an issue that needs to be more thoroughly studied. The aim of this thesis research was to investigate the population genetic structure of *Polyprion* in the Southern Hemisphere and determine whether there is a significant level of genetic differentiation between the Atlantic, Indian and Pacific ocean basins. The two specific objectives were to: 1) determine the levels of genetic variation and differentiation among *Polyprion* collection sites from Australia and New Zealand, and 2) test whether the taxonomy of *Polyprion* was supported by genetic data. To do this, DNA sequences of the mitochondrial DNA (mtDNA) control region and genotypes from nine microsatellite DNA (msatDNA) loci were used.

***Polyprion oxygeneios***

A total of 613 *P. oxygeneios* DNA sequences from the mtDNA control region and 406 genotypes from eight msatDNA loci were collected and analysed in this study. Most of the samples were from Western Australia (four sites) and New Zealand (eight sites), and eight samples were from Cape Town in South Africa. The final

mtDNA sequences were 469 base pairs (bp) long and there were 56 polymorphic nucleotide positions and 85 haplotypes. The $\Phi_{ST}$ estimates for levels of differentiation in the mtDNA data detected a significant genetic difference between Hokitika, and Chatham Islands, New Zealand, and many of the other sites in Australasia. The differences between Cape Town and most of the Australasian sites were high, but none of the estimates were statistically significant. This result could have been due to a small sample size of only eight sequences from Cape Town. Arranging the sample sites as a hierarchical set of four groups, comprising Hokitika, Chatham Islands, other Australasian sites, and Cape Town, better explained the patterns of the data set and the between group variation than any other groupings tested. In contrast, the msatDNA best supported a two-population model, with Australasia grouped as one population using the AMOVA analyses despite a significant level of differentiation detected between Hokitika and other sites in Australasia in the $\Phi_{ST}$ analyses.

***Polyprion americanus***

In total, 205 *P. americanus* DNA sequences from the mtDNA control region were analysed, half of which were sampled from Western Australia. Out of the 198 *P. americanus* samples successfully genotyped, 176 were genotyped at eight or nine microsatellite DNA loci, however the 22 samples from Argentina were successfully genotyped for six loci only. The analyses consisted of sites in Western Australian (four sites), eastern Australia (two sites), as well as 23 samples from one site in New Zealand. The final mtDNA sequences were 495 bp long, and consisted of 56 polymorphic sites and 67 haplotypes. The $\Phi_{ST}$ estimates for levels of differentiation in the mtDNA data showed that apart from the strong differentiation between Argentina and Australasia, there is very little differentiation between *P. americanus* among Australian and New Zealand sites. South West Capes (Western Australia) was however significantly differentiated from New Zealand and Taupo

Banks (eastern Australia) in the mtDNA $\Phi_{ST}$ analyses. The FST analyses of the microsatellite DNA data supported one all-inclusive Australasian grouping.

The results from this study suggested that although differentiation can be observed between ocean basins, both *P. oxygeneios* and *P. americanus* are largely homogenous at national levels and within ocean basins. Significant levels of differentiation were found between pairwise comparisons and AMOVA using samples from the Cape Town, Argentina and Australasian sites, which lends support to the suggestion of a third *Polyprion* species off the coast of South Africa and the revival of *P. moene* in New Zealand.

# *Acknowledgements*

I would like to express my gratitude to my supervisors Peter Ritchie and Corey Wakefield for their useful comments, remarks and engagement through the learning process of this master thesis. I would especially like to thank the numerous scientists and fishermen including, Sean Hennessey, Mark Green, Wayne Sumpton, and Alan Williams, who contributed extensively to the sample collection and without whom this thesis would not have been possible.

And, of course, my family. The most important people of the lot. Mum wouldn't be too impressed if I forgot to give her a special mention.

Thank you all.

# Contents

**Abstract** **i**

**Acknowledgements** **iv**

**Contents** **v**

**List of Figures** **vii**

**List of Tables** **ix**

**1 Introduction** **1**
1.1 Biology . . . . . 1
1.2 New Zealand and Australian fisheries . . . . . 3
1.3 *Polyprion* fisheries . . . . . 4
1.4 Fisheries management and the stock concept . . . . . 6
1.5 Genetic markers in fisheries . . . . . 9
1.6 *Polyprion* genetics . . . . . 11
1.7 Management implications . . . . . 12
1.8 Aims and objectives . . . . . 12

**2 The phylogeographic structure and genetic differentiation of *Polyprion oxygeneios* based on mitochondrial DNA control region sequences and microsatellite DNA genotyping** **13**
2.1 Introduction . . . . . 14
2.2 Methods . . . . . 16
2.2.1 Sample collection and DNA sequencing . . . . . 16
2.2.2 Mitochondrial DNA extraction, amplification and sequencing 18
2.2.3 Microsatellite amplification and genotyping . . . . . 18
2.3 Genetic analyses . . . . . 19
2.3.1 Genetic diversity . . . . . 19
2.3.2 Population structure . . . . . 21
2.3.3 Demographic history . . . . . 22
2.3.4 Mutation-drift equilibrium and bottleneck . . . . . 24
2.4 Results . . . . . 25

2.4.1 Genetic diversity . . . 25
2.4.2 Population structure . . . 30
2.4.3 Demographic history . . . 39
2.5 Discussion . . . 44
2.5.1 Genetic diversity and structure . . . 44

**3 The phylogeographic structure and genetic differentiation of *Polyprion americanus* based on mitochondrial DNA control region sequences and microsatellite DNA genotyping 50**
3.1 Introduction . . . 51
3.2 Methods . . . 53
3.2.1 Sample collection . . . 53
3.2.2 Mitochondrial DNA extraction, amplification and sequencing 54
3.2.3 Microsatellite amplification and genotyping . . . 55
3.3 Genetic analyses . . . 56
3.3.1 Genetic diversity . . . 56
3.3.2 Population structure . . . 58
3.3.3 Demographic history . . . 60
3.3.4 Mutation-drift equilibrium and bottleneck . . . 61
3.4 Results . . . 61
3.4.1 Genetic diversity . . . 61
3.4.2 Population structure . . . 67
3.4.3 Demographic history . . . 69
3.5 Discussion . . . 78
3.5.1 Genetic diversity . . . 78

**4 Discussion 84**
4.1 General . . . 85
4.1.1 Genetic diversity and population structure . . . 85
4.1.2 *Polyprion* taxonomy . . . 88
4.1.3 Demographics . . . 88
4.1.4 Management . . . 89
4.2 Summary . . . 90

**A Additional information for *P. oxygeneios* 92**

**B Additional information for *P. americanus* 121**

**References 137**

# List of Figures


2.1 Sampling locations and haplotype distribution for *P. oxygeneios* . . 26
2.2 Haplotype discovery curve . . . . . 28
2.3 Haplotype genealogy of the mtDNA control region sequences. Each colour represents a unique haplotype. The scale shown indicates the frequency of each haplotype. Small, black circles represent putative mutational steps between haplotypes . . . . . 29
2.4 Infinite alleles model displaying one candidate for balancing selection; PAM035, and two candidates for positive selection; GH0OIK, GJSLB2 . . . . . 30
2.5 Stepwise model displaying three candidates for positive selection; pam017, GH0OIK, GJSLB2 . . . . . 31
2.6 The linear relationship between distance and $\phi$st in *P.oxygeneios* . 34
2.7 The linear relationship between distance and $\phi$st in *P.oxygeneios* when Cape Town is excluded from the data . . . . . 34
2.8 Results for the Bayesian structure analysis of *P. oxygeneios* assuming k=3 genetic clusters. Each line represents an individual. Individuals are ordered by site . . . . . 38
2.9 Factorial components analysis (FCA) of *P. oxygeneios* showing axes 1, 2, and 3. All grey individuals in the top right corner are from HOK. All blue individuals clustered together in the middle are from CAP. All other individuals cluster together in the top left corner. . 38
2.10 All sites, Mismatch distribution of all pairwise differences. Freq. expected is under a model of population expansion . . . . . 41
2.11 Mismatch distribution of all pairwise differences for NZL. Freq. expected is under a model of population expansion . . . . . 41
2.12 Mismatch distribution of all pairwise differences for AUS. Freq. expected is under a model of population expansion . . . . . 42
2.13 Bayesian skyline plot with $\mu$=2.0x10$^{-8}$. Time=time before present. Y axis=effective population size . . . . . 42
2.14 Bayesian skyline plot with $\mu$=3.6x10$^{-8}$. Time=time before present. Y axis=effective population size . . . . . 43

3.1 Sampling locations and haplotype distribution for *P.americanus* . . 63
3.2 Haplotype discovery curve. . . . . 64

3.3 Haplotype genealogy of the mtDNA control region sequences. Each colour represents a unique haplotype. The scale shown indicates the frequency of each haplotype. Small, black circles represent putative mutational steps between haplotypes . . . . . 64
3.4 Correlation between $\Phi_{ST}$ and geographic distance . . . . . 66
3.5 Correlation between $\Phi_{ST}$ and geographic distance, excluding ARG . 66
3.6 Fst outlier results under the Stepwise Mutation Model (SMM). Both GJLSB2 and PAM035 are candidates for positive selection with He=0.289 and 0.650 respectively, and Fst=0.774 and 0.288 respectively . . . . . 67
3.7 Fst outlier results under the Infinate Alleles Model (IAM). GJLSB2 is a candidate for positive selection with He=0.289 and Fst=0.774 and 0.288 . . . . . 67
3.8 Results for the Bayesian structure analysis of P.americanus assuming k=2 genetic clusters. Each line represents an individual. Individuals are ordered by site. . . . . 71
3.9 Factorial components analysis (FCA) of *P.americanus*. The outlier in top right of the graph is an ARG individual. . . . . 72
3.10 Factorial components analysis (FCA) of *P.americanus* excluding ARG samples. . . . . 72
3.11 Mismatch distribution of all pairwise differences for all sites. Freq. expected is under a model of population expansion. . . . . 73
3.12 Mismatch distribution of all pairwise differences for Australasia. Freq. expected is under a model of population expansion. . . . . 73
3.13 Mismatch distribution of all pairwise differences for Argentina. Freq. expected is under a model of population expansion. . . . . 76
3.14 Bayesian skyline plot with $\mu$=2.0x10$^{-8}$. Time=time before present. Y axis=effective population size. . . . . 76
3.15 Bayesian skyline plot with $\mu$=3.6x10$^{-8}$. Time=time before present. Y axis=effective population size. . . . . 77

# List of Tables


2.1 *P.oxygeneios* sample locations and sample sizes . . . . . 17
2.2 New Zealand sample locations and sample sizes for *P.oxygeneios* . . 17
2.3 Nine microsatellite DNA markers and type of repeat. . . . . 19
2.4 Summary statistics for the *P. oxygeneios* mtDNA sequences. . . . . 27
2.5 Summary statistics for nine microsatellite loci. . . . . 30
2.6 Pairwise $\phi$ST values below the diagonal, with P<0.05 in bold. All values remained significant after Bonferroni correction. P-values are provided above the diagonal . . . . . 32
2.7 mtDNA results from analysis of molecular variance (AMOVA) of P.oxygeneios. . . . . 33
2.8 Pairwise FST ($\theta$) values below the diagonal, significant values after bonferroni correction indicated with * . . . . . 36
2.9 Analysis of Molecular Variance (AMOVA) results for four groupings: Region 1)CAP and AUSTRALASIA (AUS & NZL); Region 2) CAP, AUSTRALASIA (EXCL. HOK), and HOK; Region 3)CAP, AUS and NZL; and Region 4) CAP, AUS separated into WA and EA, and NZL. The % variation attributable to each model is highlighted in bold . . . . . 37
2.10 Estimation of population expansion using Fu's Fs, and mismatch distribution analyses: sum of squared deviation (SSD), Harpending's raggedness index,Tau ($\tau$) , population size before expansion ($\theta$o), population size after expansion ($\theta$1), and time since population expansion occurred (t) with P-values and confidence intervals in brackets. . . . . 40

3.1 *Polyprion americanus* samples . . . . . 54
3.2 Nine microsatellite DNA markers and type of repeat . . . . . 56
3.3 Summary statistics for *P.americanus* . . . . . 62
3.4 Summary statistics for nine microsatellite loci. . . . . 65
3.5 Pairwise $\Phi_{ST}$ values below the diagonal, with P<0.05 in bold. . . . 68
3.6 Analysis of molecular variance (AMOVA) of *P.americanus*. . . . . 70
3.7 Pairwise FST ($\theta$) values below the diagonal, significant values after bonferroni correction indicated with * . . . . . 71
3.8 Microsatellite DNA: Analysis of Molecular Variance (AMOVA) results . . . . . 74

3.9 Estimation of population expansion using Fu's Fs, and mismatch distribution analyses: sum of squared deviation (SSD), Harpending's raggedness index,Tau ($\tau$) , population size before expansion ($\theta$o), population size after expansion ($\theta$1), and time since population expansion occurred (t) with P-values and confidence intervals in brackets. . . . . 75

# Chapter 1

# Introduction

## 1.1 Biology

The genus *Polyprion*, commonly known as Wreckfish, is currently represented by two species. The bass groper *P. americanus* is widely distributed in both the Northern and Southern Hemisphere, whereas the distribution of hapuku *P. oxygeneios* is limited to continental, insular and seamount slopes of the Southern Hemisphere (Roberts 1986). Both species have an anti-tropical distribution and can be characterised as having a long Pelagic juvenile stage, only settling and becoming demersal at $>$ 50 cm total length (Wakefield et al. 2010, Wakefield et al. 2013). They belong to the family Polyprionidae, which is polyphyletic and represented by two genera, each with two species. In the Southern Hemisphere, *P. americanus* support fisheries off the Juan Fernandez Islands, Chile, New Zealand and Australia, and are generally found at depths of 40 to 800m (Francis et al. 1999, Sedberry et al. 1999).

It is thought that a long pelagic juvenile stage, coupled with an association with free-floating objects (flotsams), promotes pan-oceanic mixing of *P. americanus* populations (Wakefield et al. 2010). In the Northern Atlantic, *P. americanus* form large spawning aggregations off South Carolina, further facilitating panmixia by increasing the genetic randomness in which individuals from across the Atlantic mate with each other. *P. oxygeneios* are assumed to also form spawning aggregations in the Southern Hemisphere and experience pan oceanic mixing. Although spawning locations are unknown for both species in the Southern Hemisphere, Francis et al. (1999) speculate that Cook Strait in New Zealand could be a possible

spawning site for *P. oxygeneios*. Trawl survey data indicates the Stewart/Snares shelf and Chatham Rise in New Zealand to be important nursery sites (Paul 2002).

The need for taxonomic clarification of *P. americanus* has previously been cited (Ball et al. 2000). Using meristic and morphometric data, Robert (1986) concluded there to be two species of *Polyprion*; a deep bodied, uniform coloured *P. americanus*, and a shallow bodied, counter shaded *P. oxygeneios*. Meristic data included four internal counts; epipleural ribs, dorsal trisegmental pterygiophores, dorsal procurrent candal rays, and ventral procurrent candal rays, as well as two scale counts; tubed scales in the lateral line, and scales in oblique series between the lateral line. Although none of the meristic counts were able to reject Robert's hypothesis of a presence of two species of *Polyprion*, none of the six counts were diagnostic. Instead, three of the eight morphometric measurements directly supported his hypothesis, these being; body depth, pectoral fin length, and size of body scales. Length of longest dorsal spine, lower jaw length, and snout length were non-diagnostic, and although head depth and eye size were largely different between the two types, the overlapping values between the two left neither diagnostic. For the most part, Robert found adult *P. americanus* to be indistinguishable from each other over a wide geographic range ( i.e. the Atlantic, Indian and Pacific Oceans), although he noted four *P. americanus* juveniles from South Africa with a distinct spotted body colouration. In contrast to the mottled and banded colouring in other *P. americanus*, this one character was interpreted as plesiomorphic, as opposed to an introgression of white patterns on the body of banded and mottled individuals, which appeared to be homologous and to have occurred after the evolution of 'spots', a 'primitive' character. Thus, using the evolutionary species concept, Robert did not define these four South African individuals as separate species.

The potential for a third species of *Polyprion* has been noted by Ball et al. (2000) who recorded distinctive mitochondrial DNA (mtDNA) profiles and microsatellite genotypes in samples from South Africa. There is also speculation that *P. americanus* in New Zealand may be a separate species to *P. americanus* in the southern Atlantic, and was previously referred to as *P.moene* (Ball et al. 2000). Recent studies using molecular techniques have revealed cryptic speciation in the Southern Hemisphere where meristic and morphometric relationships have not (e.g. Choat et al. 2012, Iwatsuki 2013, Wakefield et al. 2014, Andrews et al. in press). For example, Andrews et al. (in press), has recently determined, using morphological

and molecular data, that *Etelis carbunculus* in the Indo-Pacific comprises of two reciprocally monophyletic species; *E.carbunculus*, predominantly found in the Pacific Ocean, and *E.sp.*, most common in the Indian Ocean. *E. sp.* likely diverged from *E.carbunculus* in allopatry, facilitated by phylogeographic barriers formed during the Pliocene.

## 1.2 New Zealand and Australian fisheries

New Zealand's Exclusive Economic Zone (EEZ) of 4.4 million km$^2$ is one of the largest in the world, supporting over 100 different fisheries, and contributing an approximate $1.5 billion a year to New Zealand's economy (Ministry for Primary Industries 2012). The Quota Management System (QMS) was introduced in 1986 as a principle tool to ensure long term sustainability of New Zealand's fisheries resources and counter the effects of a recent development in more intensive fishing methods. The stock concept lies at the core of the QMS, with political, commercial, and biological considerations defining the boundaries of each fishing unit/stock. Each year, major fish stocks are assessed and, at least notionally, the total allowable catch (TAC) of each stock should be set according to the number of fish that can be removed from each given stock without detriment. From the total allowable commercial catch (TACC), an annual catch entitlement (ACE) is allocated. Like other forms of property, ACE may be sold, leased, bought, or transferred. Setting ACE based on a maximum sustainable yield (MSY) requires a description of the fishery, catch records, knowledge on the species' life history characteristics, as well as biomass trends. The outcome of the assessment often reflects how well correlated the stock boundaries are with population boundaries. An OECD report (1996) found that of 37 stocks in 11 countries managed by individual quotas, initial TACs appeared to have been set too high for 24 of these, resulting in temporary stock declines and collapses (http://www.fao.org/docrep/003/W7292E/w7292e04.htm).

Australian fisheries contribute a net nominal value of AU$1.3 billion to the Australian economy, with 49.5% of this value consisting of the western rock lobster fishery alone. Being a confederation of states, fisheries legislation and policy in Australia is controlled by both state and national interests, complicating management and stock assessments. In 2002, the Ecologically Sustainable Development

(ESD) policy was put in place by the Department of Fisheries in Western Australia. The policy focuses on an ecological approach, with indicator species used to determine the health of the various bioregions and fisheries. In this case, an indicator species is "a species, for which the status of its fished stock(s) can be used as a robust indicator of the sustainability status and risks within a suite." Using indicator species is meant to ease the difficulties managers face monitoring hundreds of fish species being targeted by more than one interest. The designation as an indicator species is based on longevity, reproductive characteristics, recruitment patterns, survival after interactions with fisheries, and stock structure. *P. oxygeneios* is listed as an indicator species for the West Coast Demersal scalefish fishery, and is listed as medium high risk.

## 1.3 *Polyprion* fisheries

Throughout it's distribution, *Polyprion* are caught both recreationally and commercially. The high level of fishing effort required, coupled with *Polyprion's* mild taste and firm texture, has attracted recreational fishers and driven up demand and price. So although the Hapuku-Bass (HPB) fishery is relatively small in tonnage, the meat has a high value per kilo when compared to other species. *P. americanus* are long-lived, slow growing, and late to mature, these being life history characteristics that reflect an inherent vulnerability to over exploitation (Lane et al. 2016, Wakefield et al. 2013). In the Southern Atlantic, overfishing of *P. americanus* has resulted in a moratorium on their capture since 2005 in Brazil (Peres & Kippel 2003, Peres 2009), and full protection has been granted in South Africa following declines in landings (Heemstra 2004). There is little data available for *P. oxygeneios* in the Southern Atlantic, however this is unsurprising given that *P. oxygeneios* and *P. americanus* appear to be geographically partitioned, with *P. oxygeneios* sighted more commonly in the South Pacific (Wakefield et al. 2013).

In New Zealand, the genus *P. americanus* is managed as 'groper', a series of paired-species fish stocks (Paul 2002). This has removed an incentive for fishermen to record their catch by species, inhibiting previous attempts to effectively evaluate the total allowable catch (TAC) of both *Polyprion* species individually. Despite continued low landings, total allowable commercial catch (TACC) continually increased from 1830 t in 1986 to 2179 t in 2003 based on appeals by individual

fishers (Paul 2000). A long pelagic juvenile stage, coupled with little understanding of *P. americanus* biology in the Southern Hemisphere, leaves other forms of stocks analyses, such as otolith microchemistry, largely redundant (Paul 2002). Mark-recapture studies are also difficult due to the barotrauma experienced by deep-water species. The QMS stock boundaries were set in 1986 purely for administrative convenience, and differ from the boundaries used for determining regional yields (Paul 2002). Further, each fishery stock had it's quota set in 1986 by equivalent percentages in all areas, despite there being no catch history justification for this decision (Paul 2002)

In contrast to New Zealand's highly productive waters, south-western Australia is heavily influenced by the pole-ward flowing Leeuwin current, and can be characterised as having warmer, low nutrient, and therefore less productive waters (Molony et al. 2011). While in New Zealand, *Polyprion* spp. represent a relatively large fishery, with an annual catch of ∼1500 tonnes (t), catches of these two species off the Western Australian coast are small in comparison ($<$ 25 t, Wakefield et al. 2010, Fletcher & Santoro 2014), in part due to the nutrient poor waters in the area (Malony et al. 2011). The majority of this catch is taken by the South Coast Demersal Line Fishery. There are no output controls in place for the species, however current input controls include a maximum allowable number of lines and hooks, restricted entry, and total fishing time allocations. *P. oxygeneios* is also closed to recreational fishing from 15th October through to 15th December in the west coast bioregion.

Although the current risk to sustainability is deemed moderate due to the low catches of *P. oxygeneios* in Western Australia, recent research suggests that catches of this species could exceed levels of exploitation. Wakefield et al. (2010) suggest that although the instantaneous rate of fishing mortality (F) of *P. oxygeneios* in 2005-06 was around acceptable levels (F=0.01-0.05, Ftarget=0.06) in south-western Australia, the exploitation reference levels were derived from a lower estimate of longevity in WA compared to that determined from New Zealand (i.e. 52 v. 63 years, Francis et al. 1999). The current stock status of *P. americanus* in south-western Australia is even more uncertain, considering catch and effort statistics likely provide an unreliable index of relative abundance. This species is rare in catches, making sufficient samples sizes for representative age structures impracticable to collect. In addition, this larger congener has a higher inherent vulnerability to exploitation based on a greater longevity (52 years for *P. oxygeneios*

v. 78 years for *P. americanus*), later maturation (7 years for *P. oxygeneios* v. 11-14 years for *P. americanus*) and lower natural mortality (0.09 for *P. oxygeneios* v. 0.05 for *P. americanus*, Wakefield et al. 2010, 2013). The stock structure of these species in south-western Australia is also unknown with both commonwealth and state-managed fisheries likely exploiting the same stocks. Thus, there remains a high level of uncertainty around the current status of stocks of *Polyprion* species in south-western Australia (Wakefield et al. 2010).

## 1.4 Fisheries management and the stock concept

The main focus of fisheries management is to protect fishery resources from unsustainable levels of exploitation whilst maximising long-term economic output. Realising that all species and populations have a unique and varied response to exploitation, the stock concept was created to better manage resources (Carvalho & Hauser 1995). The idea is that an independent stock with its own sustainable yield (SY) can be designated to a set area with defined boundaries in order to reduce negative impacts on the area's long term fishing viability. It is the unique and varied responses to exploitation from each stock that underpins research and stock assessments (Carvalho & Hauser 1995). The sustainable yield will be determined by life-history strategies (e.g. longevity, maturity, growth) and population demographics (e.g. fishing and natural mortality, migration rates, population size).

Maintaining genetic variation within a population is important for its long-term persistence and adaptability. First, heterozygosity reduces the detrimental effects of inbreeding, such as inbreeding depression. Secondly, the more variation there is present within a population, the greater the chance there is of unique, advantageous alleles to help those populations adapt to a changing environment. New alleles are brought about by mutations which are either lost, or become fixed in a population through selection and random genetic drift. The creation of advantageous alleles, and genetic drift randomly fixing alleles within a population changes the frequency of alleles over long periods of evolutionary time, ultimately resulting in population divergence. Alternatively, migration (i.e. the movement of genes from one population to another) acts to homogenize populations and maintain population integrity. It is the balance between these opposing forces which results in the level of gene flow and connectivity between areas that we see today (Hedgecock et al. 2007).

Fishery managers may define a stock according to fishing dynamics within a specific area, however such definitions often fail to incorporate the actual genetic and demographic sub structuring of a species (Carvalho & Hauser 1995). Aligning fishery stocks with the substructuring of a species is made more difficult by a lack of concordance among fishery scientists and population geneticists as to what constitutes a population (Carvalho & Hauser 1995, Waples 1998, Ovenden 2013). Although many agree that spatial and temporal integrity is implicit in defining a population, structuring is continuous, making integrity difficult to measure (Lowe & Allendorf 2010). Even with a seemingly inconsequential level of migration, there may be no noticeable genetic differentiation between samples (Laikre et al. 2005, Waples 1998). Demographic independence may be a more useful definition for establishing population boundaries, however widely used techniques are unable to quantify the degree to which gene flow affects a population (Lowe & Allendorf 2010). Waples & Gaggiotti (2006) suggest that the shift from demographic dependence to independence occurs when recruits consist of less than 10% migrants. In other words, although demographic connectivity can not be quantified, it can be inferred. Between genetic and demographic independence, establishing population boundaries is more difficult and the connectivity between two sites can be labelled as 'crinkled' (Ovenden 2013).

Connectivity was, until recently, presumed to be widespread in the marine environment. However, where physical and oceanographic processes were formerly thought to be the only barriers to gene flow (Hedgecock et al. 2007), evidence suggests homing and self replenishment processes to be more common than previously thought (Cowen 2006, see Hauser & Carvalho 2008 for contrary examples).

There are a number of different ways for detecting gene flow, including non-genetic techniques (such as physical and environmental/chemical tags), direct genetic techniques (not dissimilar to non-genetic techniques), and indirect genetic techniques.

Direct techniques assign individuals to source populations, however in using this technique, it is assumed that all potential source populations are known, and that they do not depart from linkage equilibrium or Hardy-Weinberg (Hauser & Seeb 2008, Hedgecock et al. 2007). Nevertheless, this approach is useful in describing contemporary patterns of gene flow. With respect to illegal fishing and effective fisheries management, direct techniques may also increase the traceability of fish and fish products by assigning these products to their population, species, or region of origin (Hauser & Seeb 2008, Ward 2000). However, direct techniques

are of limited use in phylogeographic studies, or in any population genetic studies examining historical gene flow or events. The markers used, such as microsatellite DNA markers, are prone to homoplasy as a consequence of their high mutation rates. In other words, electromorphs (DNA fragments) of the same size may not be identical in descent, rather identical only due to convergent mutations (Estoup 2002).

Indirect methods estimate variance in allele frequencies between sites using genetic markers that have presumably reached equilibrium. However, marine populations tend to be quite large, and as such, it may take hundreds of thousands of years for a population to reach equilibrium. This results in retention of historical patterns of gene flow which may differ significantly from contemporary patterns (Hedgecock et al. 2007). While indirect methods are poor estimators of contemporary genetic exchange rates, indirect and direct methods may be used together to describe historical and contemporary gene flow. It is important to note however that while intermediate time scale events can significantly affect genetic variance, current genetic measures are unable to describe gene flow between populations over extensive time periods, and therefore may not record intermediate stochastic or recurrent events which may affect connectivity. For example, we are not able to distinguish the exchange of 1000 individuals every 100 generations, due to occasional 'sweepstake' events, from 10 individuals being exchanged every generation due to migration (Carvalho & Hauser 1995, Hauser & Seeb 2008, Hedgecock et al. 2007).

When a marker is unable to detect statistically significant differentiation between two sample sites, determining the degree of connectivity between them is more difficult. While there is reason to assume gene flow between the two sites, this does not necessarily mean they are also demographically coupled. In general terms, population genetic structure can be placed into three categories; 1) Panmixia (no differentiation), in which genetic homogeneity prevails over the geographic region considered, 2) Isolation-by-distance (continuous change), in which the genetic composition changes continuously over space, and 3) Distinct populations, where gene flow is small enough to permit genetic divergence among closely related populations (Laikre et al. 2005). The average level of differentiation among populations of marine species is low (mean Fst = 0.062), which is less than the average level reported for freshwater species (mean Fst = 0.222) (Ward 2000). While such a

low value for marine species may reflect a high number of species with a panmictic population structure, the generally larger population size of marine fish makes larger sampling sizes necessary for detecting population differences. Larger populations are able to hold more mutations at a locus (genetic marker), which may skew the expected ratio of heterozygotes to homozygotes using some diversity measures. Fst is dependent on being able to detect lower levels of heterozygosity than would be expected for the number of alleles in the population. If under sampled, the detection of a significant level of differentiation between populations may be 'masked' (Henriques et al. 2014).

It is typically considered harmless to over-split a fishery stock, however harvesting a group of individuals that belong to separate populations as if they are a single group (termed a mixed-stock fishery), may result in the extinction or severe reduction in numbers of one or more subpopulations (Allendorf et al. 2008). Evolutionary changes happen at the population level, and it is the diversity within the population that provides the resources for adaptation to environmental changes (Hauser & Carvalho 2008). Over-exploitation reduces population size, which can lower allelic diversity, affecting the long-term persistence of a population and adaptability to environmental changes. It also has the ability to reduce individual fitness by reducing heterozygosity, which increases the chance of recessive negative allele expression. Further, uneven fishing pressures have the potential to alter population subdivision by reducing migration, thereby reducing the effective population size, increasing the strength of genetic drift, and eroding variation. This can also affect non-exploited fisheries populations by reducing immigration (Allendorf et al. 2008). Even if genetic diversity is unaffected, the harvesting of fish will inevitably impact population structure, with increased mortality favouring earlier maturation, and a smaller size at maturation, as has been reported in *Gadus morhua* (Atlantic Cod) (Allendorf et al. 2008). Despite the biological considerations, fishery stocks are usually divided and managed according to historical boundaries, and political or commercial interests, often irrespective of the most biologically appropriate units and population structure (Hauser & Carvalho 2008).

## 1.5 Genetic markers in fisheries

Genetic markers vary in their modes of inheritance and their modes of evolution (Hoffman et al. 2009). A marker must provide a high enough level of variation,

but not too much that almost all individuals are unique (Hellberg 2009). Although conceptually there are only a few different classes of markers, constant refinement of techniques has prompted a general shift from enzyme-based to DNA-based markers (Chauhan & Rajiv 2010, Schloetterer 2004, Ward 2000).

Allozyme (enzyme-based) variation, which reflects changes in amino-acid substitutions, was one of the first widespread genetic tools used by fishery scientists (Ward 2000). However, with the use of the polymerase chain reaction (PCR) and DNA sequencing, using protein electrophoretic variation for identifying genetic stocks is now largely redundant. Directly analysing DNA itself not only provides a greater level of detail, as these markers tend to detect more variation, but with the advent of PCR, it is easier to be more targeted with the 'type' of DNA you want to use. This includes the choice between coding and non-coding, neutral and non-neutral markers, which are all subject to different evolutionary processes. This enables different population patterns to be detected (Ward 2000).

Rapidly evolving markers are generally favoured in studies wishing to capture more recent events, and in such cases may provide greater resolution due to their higher allelic diversity. Conversely, such markers can become saturated over time, which will confound differences between populations (Hoffman et al. 2009, Selkoe & Toonen 2006). Unlike nuclear DNA, the mtDNA does not undergo recombination. As a result the entire genome evolves in a similar way. MtDNA is therefore highly useful for studying phylogeographical patterns, as well as bottlenecks and gene flow. Conversely, nuclear loci are not as strongly influenced by other regions of the genome, and because of recombination, they more often evolve independently from each other. If selection is strong, then adaptive divergence may occur within a few generations, illustrating the strength of non-neutral processes when studying contemporary patterns of connectivity (Hauser & Seeb 2008). Thus, even if panmictic populations are detected using mtDNA and other neutral markers, stronger differentiation can still arise at non-neutral loci as selection prevents homogenisation among environmentally distinct locations.

## 1.6 *Polyprion* genetics

There is a paucity of genetic research on the connectivity and structure of *Polyprion* populations. The use of microsatellite markers by Ball et al. (2000), however, indicated a separation of *P. americanus* into three distinct genetic stocks. Ball et al. (2000) used microsatellite markers to measure the level of variation among sample sites in the Mediterranean, eastern and western North Atlantic, South Atlantic, and the South Pacific. The results showed that there was one genetic unit that comprised of northern Wreckfish, which were carried by currents in a general eastward direction from spawning sites on Blake Plateau, the Azores, and the Mid-Atlantic ridge to Madeira, the Canaries and perhaps Bermuda. The remaining fish were from two southern units: a Brazilian, or south western Atlantic stock; and a Pacific stock comprising of individuals from eastern Australia and New Zealand. Notably, this study did not include *P. americanus* from the Indian Ocean.

The use of mtDNA in an earlier study by Sedberry (1996) was unable to distinguish between *P. americanus* among Southern Hemisphere sites. It is possible that the more variable microsatellite DNA loci enabled Ball et al. (2000) to detect variation at a scale finer than that provided by mtDNA markers. Roberts (1986) writes that the most appropriate mode of speciation in *Polyprion* is allopatric speciation by subdivision. As such, the rate of speciation in *Polyprion* is expected to be slow because of the widespread genetic homogenisation of populations. Even if there is a lack of contemporary gene flow between oceanic basins, differentiation at neutral loci may not yet have arisen between all populations because drift is weak in larger populations. There have been no comparable studies done on *P. oxygeneios*, however Smith & Johnson (1985), using allozyme data, were unable to detect any variation within New Zealand. In contrast, Beentjes & Francis (1999) speculated, using tag and recapture data, that there is the possibility of a distinct stock in northern New Zealand. This study was limited by small sample size and limited statistical analyses available at the time. In the most recent genetic study of *P. oxygeneios*, Lane et al. (2016) detected differentiation between Hokitika in the South Island and other sites in New Zealand using microsatellite DNA markers. There are no obvious physical or biological barriers present between this site and others (Lane et al. 2016, Ross et al. 2009).

## 1.7 Management implications

Increasing fishing exploration into deep water coupled with a higher inherent vulnerability of *Polyprion* species creates a strong need to better understand the population connectivity. *Polyprion* are relatively long-lived and have a low reproductive output (K-selected biology), rendering them more vulnerable to overexploitation than faster growing fish species with a greater reproductive output (R-selected species). This is common in deep-water fish species suggesting 1) increased susceptibility to overexploitation at relatively lower harvest levels, and 2) they may require longer recovery periods from stock collapses than shallow-water species. Yet, due to coastal fishery stocks becoming increasingly over-exploited, fishing pressure has shifted from inshore to offshore stocks (Wakefield et al. 2010). *Polyprion* are also known for their good flesh and high market price, making them particularly attractive to fisheries. Their high growth rate during their pelagic phase adds to this attractiveness in aquaculture sciences (Machias et al. 2003).

## 1.8 Aims and objectives

There is both a lack of data, and inconsistency within the limited available research on *Polyprion* stock structure and taxonomy. Without more information it is difficult to know whether the fishing stocks are well defined and therefore able to sustainably support long-term exploitation. As a result, the aim of this study was to determine whether there is a significant level of genetic differentiation between the Atlantic, Indian, and Pacific Ocean basins, with a strong focus on the Indo-Pacific (i.e. south-western Australia, eastern Australia and New Zealand). The genetic data should fulfil two major functions: 1) Support either the current taxonomy of *Polyprion*, or the proposed revision by Ball et al. (2000); and 2) Provide information on the substructure of both *Polyprion* species which can be used to inform stocks assessments in Australia and New Zealand.

# Chapter 2

# The phylogeographic structure and genetic differentiation of *Polyprion oxygeneios* based on mitochondrial DNA control region sequences and microsatellite DNA genotyping

## 2.1 Introduction

The main goal of fisheries management is to maintain exploitation at sustainable levels while maximising the economic yield of fishing (Carvalho & Hauser 1995). One approach has been the introduction of the stock concept, with the central idea that each stock has a harvestable surplus that once removed, will not have a negative impact on the effective population size, genetic diversity, and the overall viability of future fisheries (Allendorf et al. 2008, Carvalho & Hauser 1995). The harvest of one stock should theoretically have no effect on the viability of another fishery stock (Carvalho & Hauser 1995). Although stocks have traditionally been based on management and political boundaries irrespective of the biological considerations of the species, it is becoming evident that the alignment of a stock with natural population boundaries is necessary for effective management.

The marine environment has long been considered one large homogenous ecosystem containing very few barriers to gene flow (Laikre et al. 2005). Yet, differentiation is more commonly being detected in pelagic species that have previously been assumed to comprise of large, panmictic populations. Both the biological and physical characteristics of a species and its environment influence the level of marine connectivity; these include pelagic larval duration, homing behaviour, topography, temperature, salinity and and oceanic and coastal currents (Cowen et al. 2006, Hauser et al. 2008,Dammannagoda et al. 2011, Miller et al. 2011). Wennerstroem et al. (2013) reported genetic breaks of seven commercially important fish species within the Baltic Sea. Most of the patterns of differentiation were different among species, and the patterns were able to be correlated to variation in salinity tolerances, oceanographic features, and varying life and population histories. In terms of stock management, the Wennerstroem study highlights the importance of the case by case approach that is needed when informing on stock structure.

Management of biological stocks is relatively straightforward when populations are distinctly structured, however such structuring is rare, with genetic composition generally either changing continuously over space, or having no detectable differentiation. Translating structure into stock boundaries in cases where genetic composition is continuous may be difficult, however it is possible to recognise areas of relative genetic homogeneity (Laikre et al. 2005). When no significant level of genetic differentiation can be detected, it is difficult to determine how that finding

relates to the demographic structure of the stocks, or whether the genetic marker did not have the appropriate level of resolution needed to detect the patterns of variation (Schloetterer et al. 2004). Where one genetic marker detects population differentiation between a group of individuals, another genetic marker type may detect no significant level of genetic differentiation between the same individuals. This is due to the differing inheritance modes of genetic markers. For example, genetic differentiation between Australasia and South America can be detected using microsatellite DNA (msatDNA) for two species of hakes (*Macruronus novaezelandiae* and *M. magellanicus*) and two subspecies of the southern blue whiting (*Micomesistius australia pallidus* and *M. a. australis*), but not using mtDNA (Takeshima et al. 2011). The discordance between the two markers likely results from the differences in the mutation rates, with mutations accumulating faster in the msatDNA than nuclear DNA (Chauhan et al. 2010).

Molecular markers have proven invaluable in discerning population structure where other techniques are inadequate (Laikre et al. 2005). For instance, tag and recapture studies are best applied to shallow-water species. The stability of DNA over the lifetime of individuals, and that it only requires a small sample, make it an attractive option. Genetic techniques also have a wider range of usage, such as being able to determine effective population sizes, investigate the demographic history of a population, and identify each individual in a population (Kirk et al. 2011). For example, the effective population size of snapper (*Pagrus auratus*) in Tasman Bay, New Zealand, has been shown to have declined since exploitation, and is 5x smaller in magnitude than the census population size (Hauser 2002). The characteristics of molecular markers provide a more comprehensive approach when considering the different stock markers.

*Polyprion oxygeneios*, and its sole congener *Polyprion americanus*, belong to the family Polyprionidae, commonly known as giant sea bass (or wreckfish). *P. oxygeneios* is limited to parts of the Southern Hemisphere (Robert 1986), and can be characterised as having a long pelagic juvenile stage, only settling and becoming demersal at $<$ 50 cm total length. In the Southern Hemisphere, *Polyprion* support fisheries off the Juan Fernandez Islands, Chile, New Zealand and Australia, and are generally found at depths of 40 to 800m (Francis et al. 1999, Sedberry et al. 1999). Francis et al. (1999) reported *P. oxygeneios* to live in excess of 60 years in New Zealand. In comparison, Wakefield et al. (2010) recorded a lifespan of $>$35 years (females) and $>$52 years (males). Although spawning locations are

as yet unknown, it is thought that *P. oxygeneios* form spawning aggregations and experience pan-oceanic mixing like their *P. americanus* counterparts. Francis et al. (1999) speculate that Cook Strait in New Zealand could be a spawning site. Trawl survey data indicates the Stewart/Snares Shelf and the Chatham Rise in New Zealand to be important nursery sites (Paul 2002).

The aim of the research presented in this chapter was to use molecular markers to determine whether there are statistically significant levels of genetic differentiation among sampled sites in the Southern Hemisphere, and to determine the level of genetic variation at each site. The population structure was assessed against the report in Ball et al. (2000) that a third species of *Polyprion* exists off South Africa. The second aim was to use the genetic data to investigate the demographic history of *P. oxygeneios*. To achieve these aims, DNA sequences from the mtDNA control region and msatDNA genotyping data were used. The expectation was that mtDNA sequencing would provide a greater insight into the demographic history of the species (Hedgecock et al. 2007), while msatDNA would be more suitable for detecting contemporary levels of differentiation. The matrilinear inheritance pattern of mtDNA, along with a lack of recombination, gives it one quarter the effective population size compared to typical nuclear genes. A smaller effective population size and the lack of recombination of mtDNA are both characteristics that make mtDNA more sensitive to demographic changes, and useful for studying differentiation over deeper evolutionary timescales. MsatDNA is often less conserved than mtDNA, undergoes recombination, and has a higher mutation rate, making it more useful for studying contemporary differentiation.

## 2.2 Methods

### 2.2.1 Sample collection and DNA sequencing

Tissue samples of *P.oxygeneios* were collected by recreational and commercial fishers from three areas in the Southern Hemisphere: South Africa (Cape Town), south-western Australia (Perth, South West Capes, Albany, and Esperance), and south-eastern Australia (Taupo Banks). Date, Ocean or latitude/longitude data, collector (recreational or commercial), total length (TL) (mm), depth (m), and sex were recorded for most of the individuals, and those samples lacking in ocean

or latitude/longitude data were excluded from the sequencing (table 2.1). Samples of fresh muscle tissue were stored in a 5ml vial containing absolute ethanol (i.e. 99.9%). However, the ethanol was drained and replaced by dimethyl sulphoxide (DMSO) for shipping to New Zealand, as it is a less hazardous compound, with samples promptly transferred into 80% ethanol upon arrival. Samples were refrigerated at 4.7°C.

Theory suggests a sample size of 50 includes 99% of the genetic variation within a population (Smith et al. 2008). However, sample sizes from South Africa were limited by the current fishing moratorium, and most *Polyprion* samples collected from eastern Australia were *P.americanus* (table.2.1). In accordance with the thesis objectives, the genetic structure of South Africa and Australia was compared to data collected and analysed by a previous study (Lane 2015). Henry designed the mtDNA primers used in the present study. Table.2.2 lists New Zealand (NZL) sample locations and sample sizes.

TABLE 2.1: *P.oxygeneios* sample locations and sample sizes

| Sample location | Sample size (n) | Sample size included in final analyses |
|---|---|---|
| Albany (WA) | 123 | 110 |
| Esperance (WA) | 66 | 52 |
| Perth Canyon (WA) | 21 | 20 |
| SW Capes (WA) | 166 | 129 |
| Taupo bank (EA) | 2 | 0 |
| Cape town (SA) | 8 | 8 |
| Total | 386 | 319 |

WA = Western Australia, EA = eastern Australia, SA = South Africa

TABLE 2.2: New Zealand sample locations and sample sizes for *P.oxygeneios*

| Sample location | Sample size (n) | Sample size included in final analyses |
|---|---|---|
| Leigh (LEI) | 49 | 49 |
| Hawkes Bay (HAW) | 22 | 22 |
| Cook straight (COO) | 25 | 25 |
| Makara (MAK) | 21 | 21 |
| Kaikoura (KAI) | 39 | 39 |
| Otago (OTA) | 49 | 49 |
| Hokitika (HOK) | 45 | 45 |
| Chatham Island (CHA) | 44 | 44 |
| Total | 294 | 294 |

### 2.2.2 Mitochondrial DNA extraction, amplification and sequencing

Genomic DNA for the mtDNA study was extracted using a standard phenol-chloroform protocol (Sambrook et al. 1989) and PCR reactions were performed in a TGradient thermal Block (Biometra, Goettingen, Germany). All reactions consisted of a 15 $\mu$l mix of 67 mM Tris-HCL pH 8.8, 16 mM $(NH_4)_2SO_4$, 3.0 mM $MgCl_2$, 200 $\mu$M of each dNTP, 0.5 $\mu$M of each primer, 0.6 $\mu g/\mu L$ Bovine Serum Albumin (BSA), and 1 unit of BIOTAQ$^{TM}$ DNA Polymerase (Bioline). Mitochondrial DNA control region primers previously designed for *Polyprion* were the forward primer tRNA-Pro-HPB (5'-CCTACCCCTAACTCCCAAAGC) and reverse primer CCD-HPB (5'-CCCCTTGCCCCTTAGAAAGAG). The amount of template DNA used varied from 0.5$\mu$l- 1.5$\mu$l. Thermal cycling consisted of 36 cycles: 95°C for 2 minutes (mins), followed by 36 cycles of 94°C for 30 seconds (s), 61°C for 30 s and 72°C for 45 s and a final extension at 72°C for 10 mins. The resultant 550bp amplicons were electrophoresed in 1.0% agarose gel, stained with ethidium bromide, and visualised under a UV-light source. Amplicons were purified using Exo-SAP-IT as per the manufacturer's instructions before being sent to either Massey University, New Zealand, or Macrogen, Seoul, Korea, where DNA sequencing was performed using the forward primer on an ABI Genetic Analyser.

### 2.2.3 Microsatellite amplification and genotyping

Four 96-well plates containing 384 specimens were prepared and sent to AgResearch for genotyping. To ensure consistency, the samples used were from the same individuals as those for the mtDNA sequencing , and an attempt was made to have samples from each site and species equally represented in the data. However in some cases, such as with Cape Town (CAP), this was impossible. *Polyprion* samples that were successfully extracted and sequenced were preferred over samples from individuals that did not sequence as well.

Small amounts of tissue, no greater than 1.0cm long and 0.5cm wide, were removed from the tissue samples and placed in wells with enough ethanol (80%) to fully cover the tissue pieces. Once each 96-well plate was filled and the lids sealed, samples were kept at 4.7°C until shipping to AgResearch (GenomNZ) in Invermay, New Zealand. There, microsatellite genotyping was conducted on the

samples using nine pairs of microsatellite primers (table.2.3). Genomic DNA was extracted using a Chelex® procedure (Bio-Rad laboratories) and a combination of five msatDNA loci previously published by Ball et al. (2000), and four propriety loci previously developed by AgResearch for Henry Lane for use in *P. oxygeneios*, were used in a multiplex genotyping panel. The multiplex was conducted with an annealing temperature of 56°C and a MgCl concentration of 20.nM. The primer concentration ranged from 0.1uM to 0.6 uM. One of the primers in each pair was fluorescently labelled and an ABI3730 Genetic Analyser (Applied Biosystems) was used to determine the allele sizes.

TABLE 2.3: Nine microsatellite DNA markers and type of repeat.

| Marker | Nucleotide Repeat |
|---|---|
| Pam010 | Di (GT) |
| Pam017 | Di (GT) |
| Pam021 | Di (AC) |
| Pam025 | Di (CA) |
| Pam035 | Di (GT) |
| GJLKPX | Di (AC) |
| GJSLB2 | Di (AG) |
| GGOQ6A | Di (TC) |
| GH0OIK | Tri (TTG) |

## 2.3 Genetic analyses

### 2.3.1 Genetic diversity

#### Mitochondrial DNA data

The final data set consisted of 613 *P. oxygeneios* samples. Sequences were edited to just under 500bp and aligned using GENEIOUS 6.1 (Biomatters). Summary statistics, including the number of segregating sites (S), haplotypes (H), nucleotide diversity ($\Pi$), haplotype diversity (Hd), private haplotypes (Hp) and the average number of pairwise differences (K), were collected using DNASP 5 (Librado & Rozas 2009). Due to the small sample size (2), Taupo banks was left out of these analyses. Fu's F statistics, Tajima's D, and pairwise fixation index ($\Phi_{ST}$), were measured in ARLEQUIN 3.5 (Excoffier et al. 2005).

To visualise the relationship between haplotypes, along with the frequency of each haplotype, Minimum Spanning Networks were generated in POPART (www.popart-.otago.ac.nz) using 1000 iterations. Rarefaction curves were constructed for each site to check whether the diversity observed accurately reflected the diversity of the wider population. ANALYTIC RAREFACTION 1.3 software (Holland 2003) simulates the number of haplotypes expected with each increment in sample size. The point at which a plateau is reached is meant to indicate the point at which no more haplotypes will be discovered, and a true representation of population diversity.

### Microsatellite DNA data

Evidence for null alleles and score errors was assessed using MICRO-CHECKER 2.2.3 (Van Oosterhout 2004) with 1,000 iterations. Deviations from the Hardy Weinberg equilibrium (HWE) were tested for across all sites and loci using the exact test with a $10^6$ Markov chain length and $10^5$ de-memorization steps implemented in ARLEQUIN 3.5 (Excoffier 2005). A pairwise test (likelihood ratio test) of linkage disequilibrium (LD) was also carried out in ARLEQUIN 3.5 with 10,000 permutations. For both tests, significance values were estimated at 0.95 confidence after Sequential Bonferroni corrections, which were used to reduce the chances of a type I error (incorrectly rejecting our null hypothesis).

Basic diversity indices (table.2.5) calculated using Genepop 4.2 included observed and expected heterozygosity (HO and He), number of alleles per locus, and the inbreeding coefficient (Fis). Allelic richness, and number of private alleles were calculated using HP-Rare. Allele frequency distributions were then generated for each population and locus using GENALEX 6.5 (Peakall & Smouse 2006, Peakall & Smouse 2012).

An Fst outlier detection method was performed in LOSITAN to test the theory of microsatellite loci neutrality (Antao et al. 2008). This method is able to identify candidates under selection by analysing the relationship between Fst and heterozygosity (Beaumont & Nichols 1996). Theoretically, loci under selection, or loci that are linked to genes under selection, will appear as outliers when compared with those influenced by drift and gene flow only. Both a stepwise (SMM) and infinite allele (IAM) mutation model were performed using 100,000 simulations and a confidence interval of 0.995. To ascertain whether sample sizes were likely

large enough to contain a level of diversity representative of the population, allele rarefaction curve graphs were generated. First, Hp-rare was used to generate a hypothetical number of discovered alleles for cumulative sampling size, with each increment being one individual. Secondly, these 'discovered' alleles were plotted against cumulative sample sizes, with a plateau in the relationship assumed to be the point at which greater sampling will not reveal new alleles.

### 2.3.2 Population structure

#### Mitochondrial DNA

Pairwise genetic differences ($\Phi_{ST}$) were used to estimate the level of gene flow between all sites using 1,000 permutations at 0.95 confidence after Sequential Bonferroni corrections in ARLEQUIN 3.5. To determine whether a pattern of Isolation-By-Distance (IBD) is present, linearised ($\Phi_{ST}$) values were plotted against geographic distance (km) using GOOGLE EARTH 7.1. The distances were based on the shortest possible distance, by sea, between two sites. Due to the small sample size of TAU, this was again excluded from analyses. Statistical significance of the linear regression was assessed using a Mantel test conducted in ARLEQUIN 3.5 using 10,000 permutations.

Analysis of Molecular Variance (AMOVA) was used to estimate the level of genetic differentiation within and among sampling areas using ARLEQUIN 3.5. 1) CAP, Western Australia (WA), and New Zealand (NZL) 2) CAP, WA, NZL, Hokitika (HOK), and the Chatham Islands (CHA) 3) CAP, WA, NZL, CHA 4) CAP, Australasia, HOK, CHA 5) CAP, Australasia, CHA, and 6) CAP, Australasia, and HOK. These groupings were based on the results from the $\Phi_{ST}$ analyses.

#### Microsatellite DNA data

Three approaches were used to measure genetic distance and estimate levels of genetic differentiation between sites. First, Weir & Cockerham's (1984) Fst analysis was implemented using FSTAT 2.9.3.2. (Goudet 1995). Fst estimates were used without assuming Hardy Weinberg equilibrium, 1,000 iterations, 95% confidence intervals, as well as Bonferroni corrections. Then, both G'st (Hedrick 2005)

and Jost's D estimator, Dst (Jost 2008) were estimated using GENODIVE (Meirmans & Van Tienderen 2004). All three measures were used to account for the limitations of each measure on their own. Fst, while widely used, bases genetic distance on decreased heterozygosity, which, in multi-allelic data, could theoretically be high even if sites contained a completely different set of alleles. With highly polymorphic microsatellite data, this can result in a downward bias in Fst values. While both G'st and Jost's D account for variation in the maximum obtainable distance value, they are both often biased upwards instead. Jost's D is measured using the effective number of alleles rather than heterozygosity, however it is also sensitive to changes in mutation rate. Due to a lack of data at loci PAM010, PAM025, and GH0OIK in CAP, these three loci were excluded from distance measures when CAP was included in analyses. The significance of a pattern of IBD was examined using ARLEQUIN 3.5.1 by running a Mantel test with 1,000 permutations. Analysis of Molecular Variance (AMOVA) was then implemented in ARLEQUIN 3.5 using four groupings: 1) CAP and Australasia 2) WA, NZL, and HOK 3) CAP, WA, NZL, and HOK 4) CAP, Australasia, and HOK, and 5) WA, NZL, and HOK.

Using STRUCTURE 2.3.4 (Pritchard et al. 2000), a Bayesian clustering analysis was performed. Each cluster (K), which was set to a range of 1-7 and simulated 10 times, is characterised by a set of allele frequencies which each individual is then assigned to. Markov chain Monte Carlo (MCMC) simulations were run 10,000, 100,000, and 400,000 times, after a burn-in period of $10^5$ steps, under both admixture and non-admixture models with the 'locprior' function implemented. Delta K ($\Delta$K), used to estimate the value of K that best reflects the population's structure, was measured using STRUCTURE HARVESTER 0.6.94 (Earl & VonHoldt, 2012). A Factorial Component Analysis (FCA) was also conducted in GENETIX (Belkhir et al. 2000) to visually examine the differences between each individual in the sample using three main axes.

### 2.3.3 Demographic history

#### Mitochondrial DNA data

The demographic history was investigated using three approaches. First, the neutrality tests Fu's F (Fu, 1997) and Tajima's D (Tajima 1989) were calculated in

ARLEQUIN 3.5 using 1000 simulations to measure the number of segregating sites in the data against the observed nucleotide diversity. Significant negative values (P<.02) indicate an excess of segregating sites given the observed nucleotide diversity, expected in a population that has recently undergone expansion (Fu 1997; Tajima 1989).

The second approach was Mismatch distribution analysis, conducted in DNASP 5. Here, the frequency of pair-wise differences between haplotypes is measured. The demographic history of a population may be inferred by the shape of the distribution, which is expected to change under different scenarios. For example, Unimodal distribution is expected in a population that has recently undergone population expansion, whereas multimodal or erratic distribution reflects more long term stability (Rogers et al. 1992). To fit the data to a sudden expansion model (1,000 replicates), both Harpending's raggedness index (Harpending 1994) and the sum of squared deviations (SSD) were calculated using ARLEQUIN 3.5. The calculations were made both by separating CAP, AUS, NZL HOK, and ESP into five groups. Harpending's raggedness index and SSD for HOK and ESP were calculated separately to other NZL and AUS sites based on FST and AMOVA results. To examine the changes in population size, the demographic parameters $\tau$, and $\theta$ were estimated from the mismatch distribution in ARLEQUIN 3.5, and the equation t= $\tau/2\mu$ (Rogers & Harpending 1992), where $\mu$ is the mutation rate of the sequence (not per nucleotide), used to estimate a time since population expansion. The value of $\mu$ can be determined using the formula $\mu$= $2\mu$k, where $\mu$ is the mutation rate per nucleotide site and k is the number of analysed nucleotide base pairs. The mutation rates of 2.0 x $10^{-8}$ proposed by Brown (1979) for the vertebrate mitochondrial genome based on mammalian data and 3.6 x $10^{-8}$ calibrated by Donaldson et al. (1999) for snook (Percoidei; Centropomidae) were used both because no mutation rate has been calibrated for *Polyprion*, and to remain consistent with Lane et al. (2016) who also used both mutation rates.

Lastly, Bayesian skyline plots were constructed in BEAST 1.8.0 (Drummond 2012) for *P. oxygeneios* excluding CAP and HOK. This method quantifies the relationship between the genealogy of the sequences and the demographic history of the population (Ho 2011). The Marko Chain Monte Carlo simulations were run for $10^7$ iterations with the HKY substitution model, and a strict molecular clock. Two mutation rates were used: 2.0x$10^{-8}$ and 2.6 x$10^{-8}$. The results were checked, and skyline plots were constructed in TRACER (Rambaut 2014).

### 2.3.4 Mutation-drift equilibrium and bottleneck

#### Microsatellite DNA data

To determine whether samples were in mutation-drift equilibrium, a mode test was implemented using BOTTLENECK 1.2.02 (Piry et al. 1999). This method compares the proportion of alleles at low frequencies to an expected L-shaped distribution in a population at equilibrium. When population sizes decrease and the strength of genetic drift is stronger, alleles at low frequencies are eliminated from the population sooner than those at a higher frequency, skewing the distribution. Allele frequency distribution plots were created using GENALEX (Peakall & Smouse 2006, 2012). BOTTLENECK was again used to run a Wilcoxon sign-rank test, which is also based on measures of heterozygosity. Three mutation models were used for the simulation: (i) the Stepwise Mutation Model (SMM), ii) the Infinite Alleles Model (IAM), and iii) the Two Phase Model (TPM). Analyses were performed with 1000 iterations, and the TPM was set to 95% single-step mutations and 5% multi-step mutations.

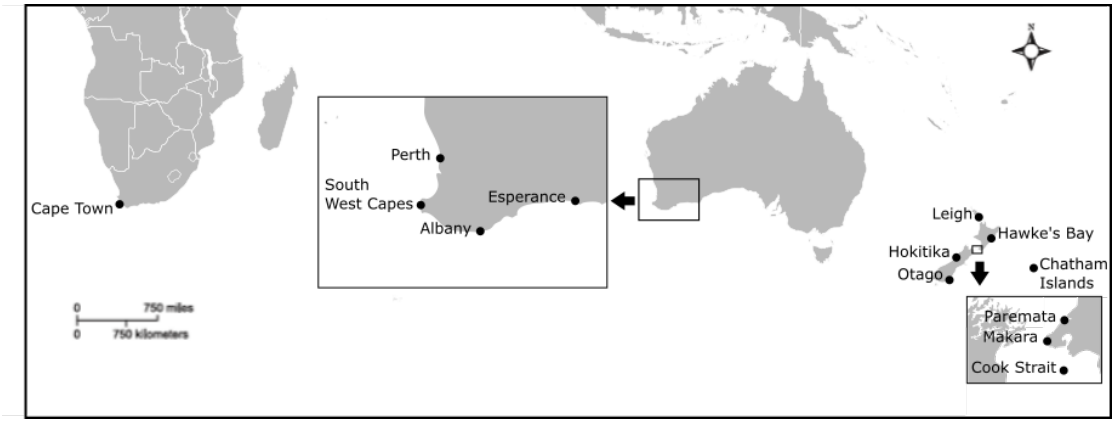
## 2.4 Results

### 2.4.1 Genetic diversity

#### Mitochondrial DNA data

In total, 613 *P.oxygeneios* DNA sequences from the mtDNA control region were analysed in this study (Table 2.4). Most samples were from Western Australia and New Zealand (311 and 217 respectively), with only eight samples having been collected from Cape Town. The sequences were aligned, and trimmed to 469bp long. The sequences showed asymmetric base frequencies of A=0.34, C=0.21, G=0.14, and T=0.31.

There were 56 Polymorphic sites, with 48 of these being parsimony informative. Forty-six polymorphic sites contained two indels, and the other two contained three indels. Hd ranged from 0.429 in Capetown to 0.957 in Hawkes Bay. In total, 85 haplotypes were present, almost half of which (44) were present in South West Capes. Haplotype 5 was the most abundant haplotype and was observed at all sites (Figure 2.1).



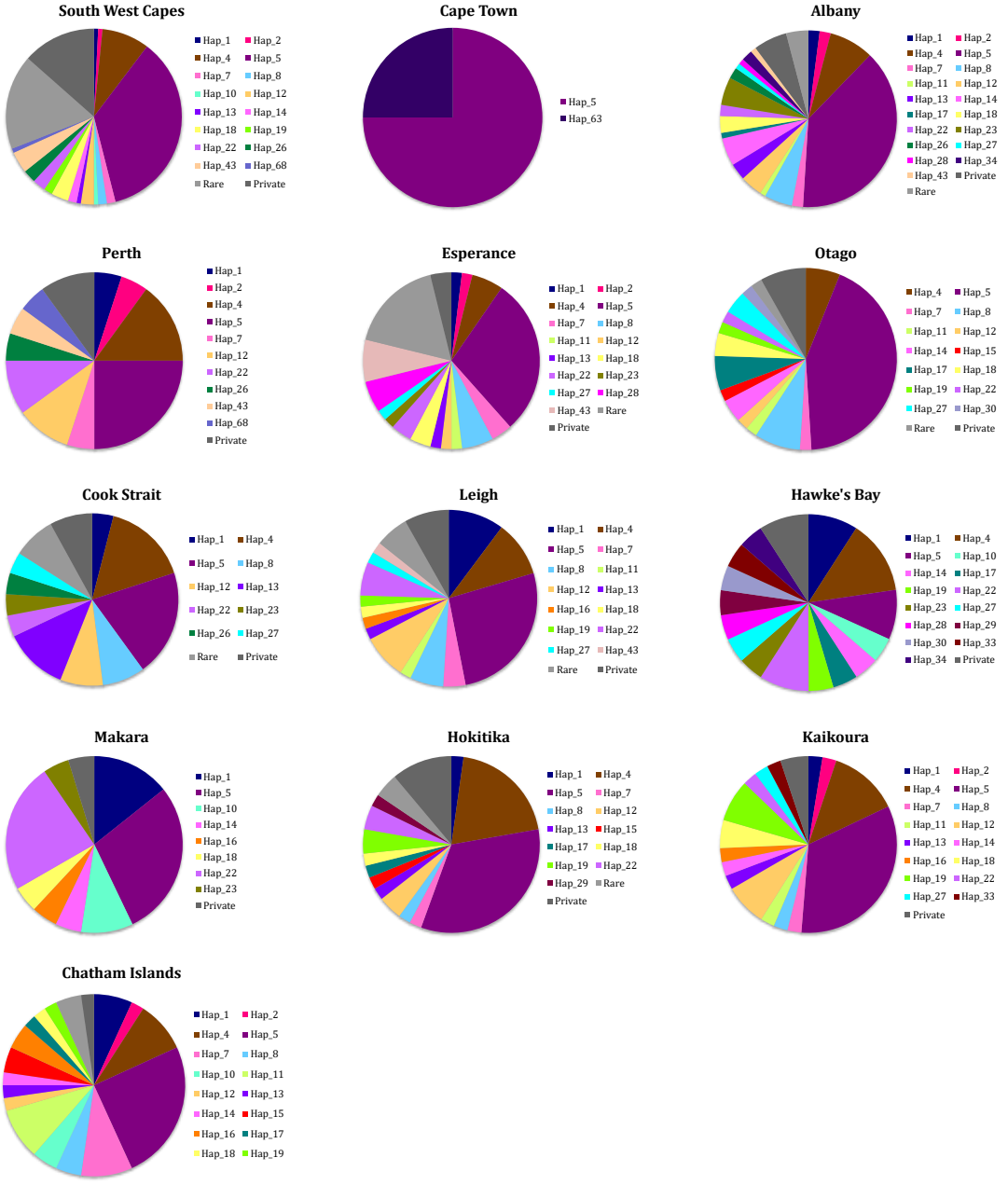




FIGURE 2.1: Sampling locations and haplotype distribution for *P. oxygeneios*

TABLE 2.4: Summary statistics for the *P. oxygeneios* mtDNA sequences.

| Location | | long | lat | N | h | Hd | S | K | Hp | $\pi$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Chatham Island | CHA | 176.81 | –43.95 | 44 | 17 | 0.87 | 30 | 3.91 | 1 | 0.008 |
| Cook Straight | COO | 174.82 | –41.79 | 25 | 13 | 0.88 | 25 | 4.63 | 2 | 0.01 |
| Hawkes Bay | HAW | 177.38 | –39.51 | 22 | 15 | 0.96 | 26 | 6.17 | 2 | 0.013 |
| Hokitika | HOK | 170.93 | –42.7 | 45 | 17 | 0.82 | 21 | 3.29 | 5 | 0.007 |
| Kaikoura | KAI | 173.74 | –42.42 | 39 | 16 | 0.85 | 27 | 4.81 | 1 | 0.01 |
| Leigh | LEI | 174.82 | –36.27 | 49 | 19 | 0.89 | 28 | 4.87 | 3 | 0.01 |
| Makara | MAK | 173.53 | –41.23 | 21 | 8 | 0.81 | 18 | 5.01 | 1 | 0.01 |
| Otago | OTA | 170.93 | –45.94 | 49 | 18 | 0.81 | 31 | 5.47 | 3 | 0.012 |
| Albany | ALB | 118.8 | –35.04 | 110 | 23 | 0.84 | 39 | 5.15 | 6 | 0.011 |
| Esperance | ESP | 122.12 | –34.5 | 52 | 22 | 0.89 | 31 | 5.71 | 2 | 0.012 |
| Perth | PER | 115.13 | –31.97 | 20 | 12 | 0.92 | 20 | 5.43 | 2 | 0.012 |
| South West Capes | SWC | 114.46 | –33.69 | 129 | 38 | 0.84 | 37 | 4.68 | 15 | 0.01 |
| Cape Town | CAP | 18.15 | -33.95 | 8 | 2 | 0.43 | 2 | 0.86 | 0 | 0.002 |

N=number of sequences, H=number of haplotypes, Hd=haplotype diversity, S=segregating sites, K=pairwise differences, Hp=private haplotypes, $\pi$=nuceleotide diversity.

The rarefaction analyses (Figure 2.2) did not reach a plateau for the most part, indicating that more haplotypes would be discovered with a greater sampling size. Interestingly, CAP appeared to reach a plateau, despite a low sample size of eight. This could reflect the difficulties in analysing data from smaller sample sizes, or it could reflect low genetic diversity at this site. It appears that the number of haplotypes seen at each site would differ, even if sample size was kept consistent throughout all sites. The remaining rarefaction curves can be viewed in Appendix A.

The most common haplotype was Hap5, followed by Hap4, as seen in the Minimum Spanning Network (Figure 2.3). The network graph showed that most haplotypes radiated out from Hap5, including Hap4, which was separated from Hap5 by only two mutational steps. The general pattern of the network was one of high haplotype diversity and in some cases haplotypes differed from each other by 10+ mutational steps. The black circles in the network represent mutational steps that were not represented by a haplotype. The high number of mutational steps that were not represented, along with the lack of a star-like shaped network, does not support a theory of sudden population expansion, rather it matches haplotype networks observed in older and more stable populations.

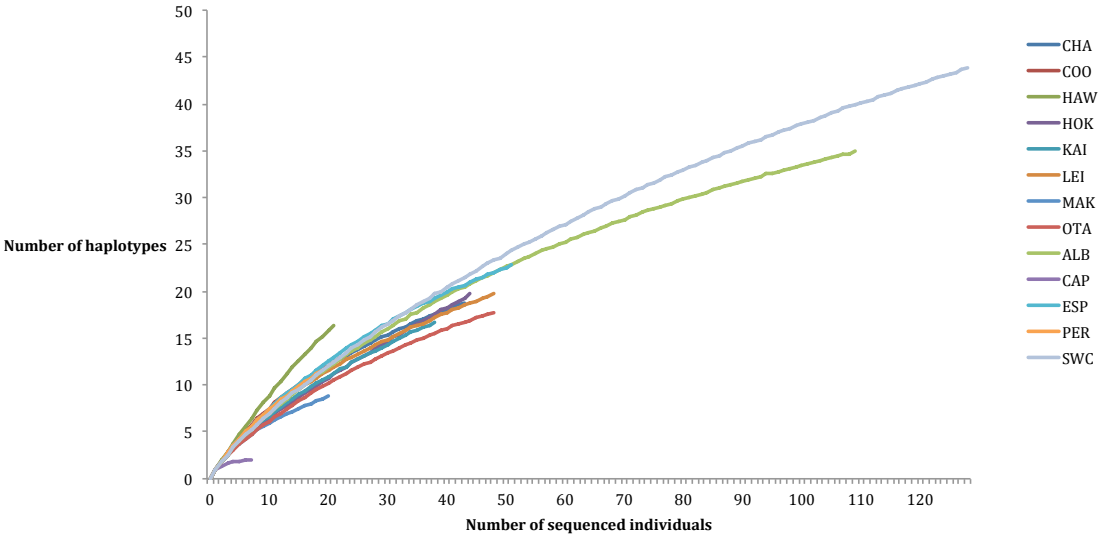



FIGURE 2.2: Haplotype discovery curve

### Microsatellite DNA data

In total, 408 *P. oxygeneios* samples were successfully genotyped (Table 2.5). Samples were removed from the final analyses if they were contaminated, or if data was missing at three or more loci. Testing for homozygote excess using Microchecker at 95% confidence provided evidence for an excess in null alleles in the sample set from HOK (n=49) at the loci GG0OQ6A, GH0OIK, GJSLB2, PAM017, and PAM025. At loci GG00Q6A and PAM025, this may be due to stuttering, however there was no evidence for stuttering at the remaining loci. There was no evidence for large allele dropout at this site. Other sites with null alleles include: ALB, at site GH0OIK (n=46, He= 14.826 Ho=21); ESP, at site PAM035 (n=27, He=5.22, Ho=12); OTA, at site PAM021 (n=49, He=17.041, Ho=22); and CHA, at site PAM035 (n=43, He=10.36, Ho=16). It is relatively common for null alleles to be present in populations that have diverged from the population for which the primers were designed (Chapuis 2007). In this case, all PAM loci were designed for use in *P.americanus* and not the conspecific *P.oxygeneios* (Ball et al. 2000). There was insufficient data available for CAP (n=7) to determine whether null alleles were present among the samples.

HOK showed a significant departure from HWE at all loci after Bonferroni correction ($P<0.00068$). Further, all loci exhibited linkage disequilibrium ($P<0.0014$). OTA showed a departure from HWE at two sites, GJSLB2 and PAM010. This data corresponded with the tests for LD, which showed OTA, similar to HOK,

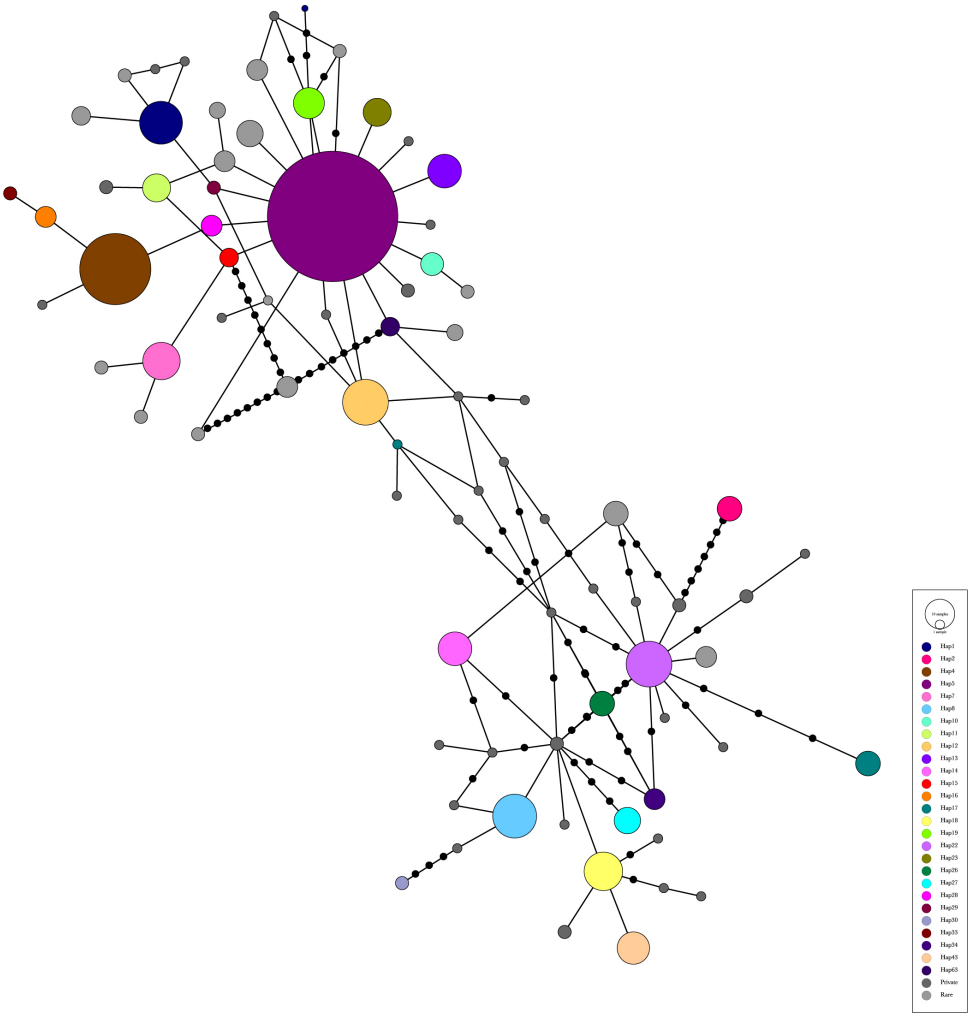


FIGURE 2.3: Haplotype genealogy of the mtDNA control region sequences. Each colour represents a unique haplotype. The scale shown indicates the frequency of each haplotype. Small, black circles represent putative mutational steps between haplotypes

to be linked at most loci. HAW also showed linkage between GG0OQ6A and PAM017, however there was no evidence for null alleles or a departure from HWE at this site.

TABLE 2.5: Summary statistics for nine microsatellite loci.

| Sample site | | N | H0 | He | Ar | PAr | Fis |
|---|---|---|---|---|---|---|---|
| Albany | ALB | 110 | 0.7 | 0.7 | 2.74 | 0.17 | 0.05 |
| Cape Town | CAP | 8 | 0.64 | 0.62 | 2.2 | 1.02 | –0.02 |
| Esperance | ESP | 52 | 0.73 | 0.78 | 2.96 | 0.75 | 0.07 |
| Perth | PER | 20 | 0.74 | 0.78 | 2.94 | 0.22 | 0.05 |
| South West Capes | SWC | 129 | 0.71 | 0.73 | 2.75 | 0.17 | 0.04 |
| Leigh | LEI | 49 | 0.77 | 0.77 | 2.87 | 0.18 | –0.009 |
| Hawkes Bay | HAW | 22 | 0.78 | 0.77 | 2.87 | 0.22 | –0.02 |
| Makara | MAK | 21 | 0.81 | 0.77 | 2.87 | 0.21 | –0.05 |
| Cook Strait | COO | 25 | 0.81 | 0.8 | 2.99 | 0.26 | –0.02 |
| Chatham | CHA | 44 | 0.73 | 0.77 | 2.89 | 0.17 | 0.06 |
| Otago | OTA | 49 | 0.76 | 0.75 | 2.81 | 0.16 | –0.02 |
| Hokitika | HOK | 45 | 0.62 | 0.81 | 3.05 | 0.55 | 0.24 |

N=Sample size (N), H0=observed heterozygosity, He=expected heterozygosity, Ar=allelic richness, PAr=Private allelic richness and Fis=fixation index.

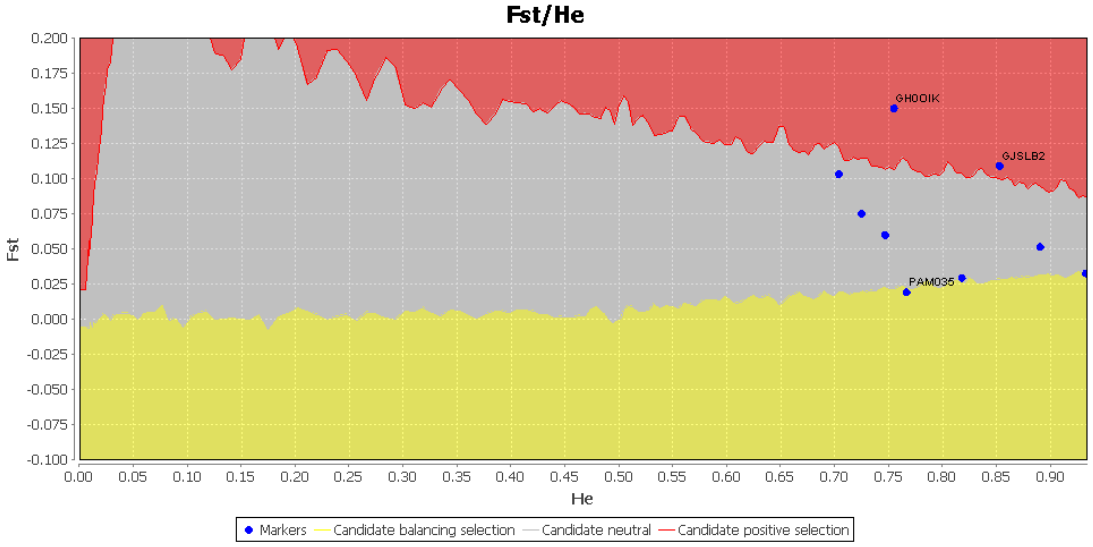



FIGURE 2.4: Infinite alleles model displaying one candidate for balancing selection; PAM035, and two candidates for positive selection; GH0OIK, GJSLB2

## 2.4.2 Population structure

### Mitochondrial DNA analysis

The $\phi$st values indicated significant differentiation between CHA and HAW, MAK, OTA, ALB, ESP, PER, and HOK and HAW, MAK, OTA, ALB, ESP, and SWC (Table 2.6). Although $\phi$st values between CAP and many other sites were relatively high, none of the values were statistically significant, perhaps due to the low

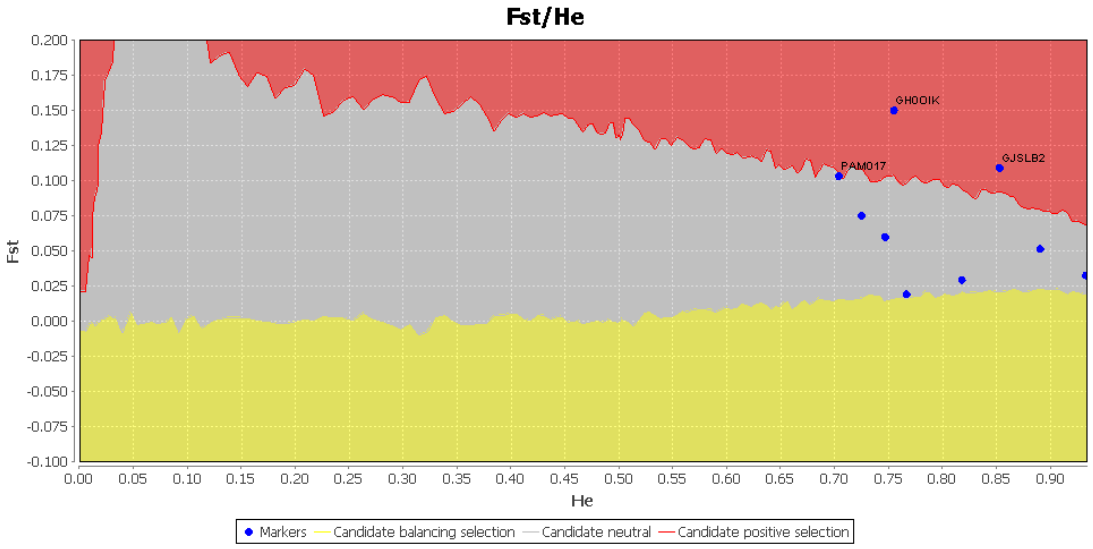



FIGURE 2.5: Stepwise model displaying three candidates for positive selection; pam017, GH0OIK, GJSLB2

sample size of CAP (8). A pattern of isolation-by-distance was observed (R=0.02) only when CAP was included in the samples. Once CAP was removed, an inverse pattern was seen, where $\phi$st increased with a decline in distance (FIGS.2.6, 2.7).

Very little differentiation was explained by among group variation (%VAR) when NZL and WA sites were grouped separately (%VAR 0.19-1.16) (Table 2.7). Although, once NZL and WA were grouped together, and HOK and CHA sites treated separate to the rest of Australasia in accordance with $\phi$st results, the level of differentiation explained by among group variation increased (%VAR 1.88-2.67) and the %VAR among sites within groups decreased (0.21-0.72). Among group %VAR was greatest using the model HOK/CHA/AUSTRALASIA/CAP (%VAR 2.67).

### Microsatellite DNA data

The pairwise FST analyses revealed significant differentiation between CAP and all other sites excluding OT, HB, and PA. Other sites that were significantly differentiated were between OA and OTA, HOK, and LEI, and between HOK and CS, CH, LEI, and OTA (TABLE 2.8). Both Nei's G'st and Jost's D results were concordant with the FST values (Appendix A). However, the AMOVA analysis was most supportive of a two-group model whereby the whole of Australasia was grouped together with CAP grouped separately (TABLE 2.9). Under this model, over 22% of the variation was attributed to between group variation. In contrast,

TABLE 2.6: Pairwise $\phi$ST values below the diagonal, with P<0.05 in bold. All values remained significant after Bonferroni correction. P-values are provided above the diagonal

| | CHA | COO | HAW | HOK | KAI | LEI | MAK | OTA | ALB | CAP | ESP | PER | SWC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CHA | | 0.38965 | **0.01367** | 0.23047 | 0.25586 | 0.16016 | **0.01172** | **0.00586** | 0.04004 | 0.31934 | **0.02246** | **0.02441** | **0.01855** |
| COO | 0.00102 | | 0.2998 | 0.49219 | 0.87207 | 0.9209 | 0.21582 | 0.09961 | 0.42676 | 0.25391 | 0.20312 | 0.3877 | 0.34668 |
| HAW | 0.07226 | 0.00547 | | **0.03125** | 0.20117 | 0.30762 | 0.9375 | 0.25098 | 0.22656 | 0.05176 | 0.33398 | 0.90039 | 0.15039 |
| HOK | 0.00499 | -0.00503 | 0.06884 | | 0.57324 | 0.15918 | **0.00879** | **0.00586** | **0.04395** | 0.33984 | 0.01074 | 0.06055 | **0.03418** |
| KAI | 0.00487 | -0.01787 | 0.01139 | -0.00721 | | 0.83984 | 0.11426 | 0.10645 | 0.47949 | 0.28516 | 0.17871 | 0.41797 | 0.45215 |
| LEI | 0.0107 | -0.02058 | 0.00364 | 0.01151 | -0.01293 | | 0.2334 | 0.12695 | 0.52734 | 0.20117 | 0.42969 | 0.57812 | 0.54004 |
| MAK | 0.08198 | 0.0148 | -0.03266 | 0.09264 | 0.0282 | 0.00804 | | 0.1875 | 0.22559 | 0.05273 | 0.35742 | 0.76953 | 0.1543 |
| OTA | 0.07006 | 0.02712 | 0.00763 | 0.07351 | 0.02177 | 0.01418 | 0.01592 | | 0.22656 | 0.08496 | 0.54492 | 0.28906 | 0.11523 |
| ALB | 0.02421 | -0.0041 | 0.00874 | 0.02314 | -0.0035 | -0.00424 | 0.00897 | 0.0049 | | 0.21289 | 0.44336 | 0.47461 | 0.65332 |
| CAP | 0.00558 | 0.01807 | 0.1009 | 0.00448 | 0.01836 | 0.03521 | 0.12583 | 0.08347 | 0.02933 | | 0.09863 | 0.06738 | 0.19141 |
| ESP | 0.04169 | 0.00971 | 0.00069 | 0.05487 | 0.00997 | -0.00305 | 0.0002 | -0.00694 | -0.00251 | 0.07002 | | 0.6416 | 0.28418 |
| PER | 0.05233 | -0.00307 | -0.02763 | 0.04473 | -0.00385 | -0.01139 | -0.02609 | 0.00476 | -0.00673 | 0.09711 | -0.01261 | | 0.53223 |
| SWC | 0.03103 | -0.00027 | 0.01484 | 0.02142 | -0.00313 | -0.00406 | 0.0151 | 0.0108 | -0.00334 | 0.02799 | 0.00206 | -0.00731 | |

TABLE 2.7: mtDNA results from analysis of molecular variance (AMOVA) of P.oxygeneios.

| Regions | Among groups | | | | | Among sites within groups | | | | | Within sites | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d.f. | Var | % Var | Fct | P-value | d.f. | Var | % Var | Fsc | P-value | d.f. | Var | % Var | Fct | |
| HOK/CHA/–Australasia/CAP | 3 | 0.077 | 2.67 | 0.02673 | 0 | 9 | 0.00607 | 0.21 | 0.00218 | 0.36 | 600 | 2.7819 | 97.12 | 0.02885 | 0.03 |
| CHA/–Australasia/CAP | 2 | 0.0693 | 2.42 | 0.024115 | 0.01 | 10 | 0.01807 | 0.63 | 0.00645 | 0.12 | 600 | 2.7819 | 96.96 | 0.03045 | 0.02 |
| HOK/–Australasia/CAP | 2 | 0.0537 | 1.88 | 0.0188 | 0.04 | 10 | 0.02067 | 0.72 | 0.00738 | 0.1 | 600 | 2.7819 | 97.4 | 0.02604 | 0.02 |
| NZL/AUS/CAP | 2 | 0.00427 | 0.15 | 0.00152 | 0.27 | 10 | 0.02792 | 0.99 | 0.00994 | 0.05 | 600 | 2.7819 | 98.86 | 0.01144 | 0.02 |
| HOK/CHA/–NZL/AUS/CAP | 4 | 0.03272 | 1.16 | 0.01159 | 0.02 | 8 | 0.00743 | 0.26 | 0.00266 | 0.34 | 600 | 2.7819 | 98.58 | 0.01423 | 0.02 |
| CHA/NZL/–AUS/CAP | 3 | -0.00534 | -0.19 | -0.00189 | 0.54 | 6 | 0.04538 | 1.61 | 0.01606 | 0.05 | 600 | 2.77954 | 98.58 | 0.0142 | 0.01 |

Six apriori geographic groupings: 1)Hokitika and Chatham Islands grouped separately to the rest of Australasia and Cape Town, 2) Chatham Islands grouped separately to the rest of Australasia and Cape Town 3) Hokitika grouped separately to the rest of Australasia and Cape Town, 4) New Zealand, Australia, and Cape Town grouped separately, 5) Hokitika and Chatham Islands grouped separately to the rest of Australasia and Cape Town, and 6) Chatham Islands grouped separately to the rest of Australasia and Cape Town . D.f. is the degrees of freedom, Var is variance component, % Var is variance component in percentage of total variation.

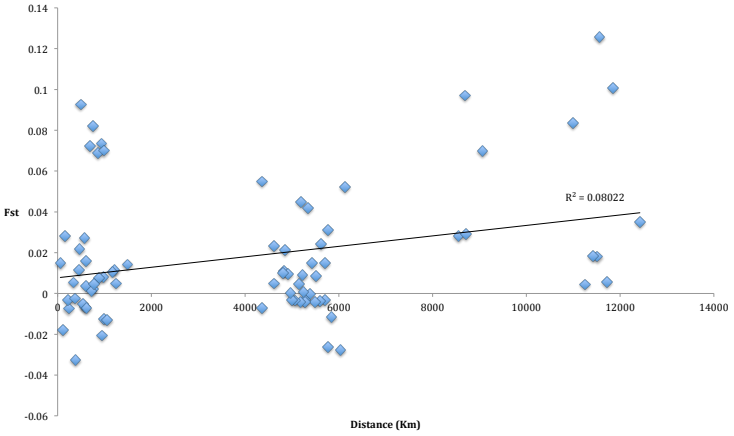



FIGURE 2.6: The linear relationship between distance and $\phi$st in *P.oxygeneios*

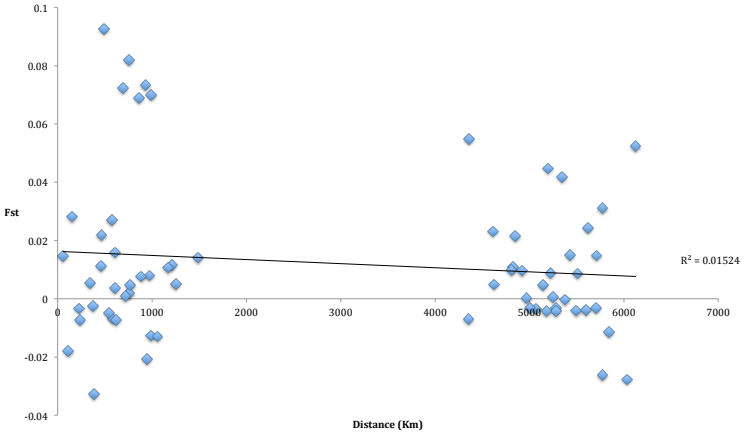



FIGURE 2.7: The linear relationship between distance and $\phi$st in *P.oxygeneios* when Cape Town is excluded from the data

8.6% of variation was attributed to between group variation under a three-group model with Australasia separated into WA, NZL, ad HOK groupings. The mantel test did not support a pattern of Isolation-by-distance either ($R^2$=-0.0008, P=0.3) Just over 2% of the variation was attributed to between group variation when HOK was separated from the rest of Australasia. The four groupings shown were chosen based on the Fst results.

While CAP individuals were grouped separately to others in the Bayesian Cluster Analysis regardless of run length, HOK individuals were grouped separately only once STRUCTURE was set to 400,000 runs, as shown in FIG.2.8. Once run length was set to 400,000, STRUCTURE supported a three-model cluster (K=3). The Factorial Components Analysis (FCA) supported the STRUCTURE Bayesian analysis results, with HOK and CAP individuals clustered together. In FIG. 2.9 all CAP individuals (blue) group together in the middle, with HOK plus one PER individuals clustered in the top right. See Appendix A for analyses excluding CAP and HOK.

Table 2.8: Pairwise FST ($\theta$) values below the diagonal, significant values after bonferroni correction indicated with *

| | OA | OC | OE | OP | OS | OT | CS | HB | MA | OTA | CH | HO | LE | PA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| OA | | * | NS | NS | NS | NS | NS | NS | NS | * | NS | * | * | NS |
| OC | 0.2873 | | NS | * | * | NS | * | NS | * | * | * | * | * | NS |
| OE | -0.00251 | 0.07002 | | NS | NS | NS | NS | NS | NS | NS | * | * | NS | NS |
| OP | -0.0003 | 0.2514 | -0.0121 | | NS | NS | NS | NS | NS | NS | NS | NS | NS | NS |
| OS | -0.0023 | 0.2865 | 0.0021 | -0.0002 | | NS | NS | NS | NS | * | NS | * | NS | NS |
| OT | 0.0075 | 0.334 | 0 | -0.0065 | 0.0087 | | NS | NS | NS | NS | NS | NS | NS | NS |
| CS | 0.0014 | 0.2632 | 0.0097 | -0.0019 | 0.0049 | 0.0045 | | NS | NS | NS | NS | * | NS | NS |
| HB | 0.002 | 0.2783 | 0.0006 | 0.0006 | 0.0043 | 0.0297 | -0.0065 | | NS | NS | NS | NS | NS | NS |
| MA | -0.0008 | 0.285 | 0.0002 | -0.006 | 0.0019 | 0.0199 | -0.0064 | -0.0088 | | NS | NS | NS | NS | NS |
| OTA | 0.0151 | 0.2767 | -0.0069 | 0.0113 | 0.0155 | 0.042 | 0.0058 | 0.0086 | 0.0068 | | NS | * | NS | NS |
| CH | 0.01 | 0.2636 | 0.0419 | -0.0011 | 0.0079 | 0.008 | -0.0025 | -0.0031 | -0.004 | 0.0046 | | * | NS | NS |
| HO | 0.0643 | 0.1855 | 0.0549 | 0.0357 | 0.0693 | 0.0347 | 0.041 | 0.0484 | 0.0459 | 0.0601 | 0.0521 | | * | NS |
| LE | 0.0086 | 0.2693 | -0.0030 | -0.0006 | 0.0058 | 0.0313 | 0.0003 | -0.0023 | -0.0031 | 0.0063 | -0.0044 | 0.0547 | | NS |
| PA | 0.0185 | 0.3375 | 0 | 0.0161 | 0.0212 | 0.0701 | 0.0233 | 0.0261 | 0.0103 | 0.0246 | 0.0181 | 0.0733 | 0.0213 | |

TABLE 2.9: Analysis of Molecular Variance (AMOVA) results for four groupings: Region 1)CAP and AUSTRALASIA (AUS & NZL); Region 2) CAP, AUSTRALASIA (EXCL. HOK), and HOK; Region 3)CAP, AUS and NZL; and Region 4) CAP, AUS separated into WA and EA, and NZL. The % variation attributable to each model is highlighted in bold

| Regions | Among groups | | | | | Among sites within groups | | | | | Within sites | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | d.f. | Var | % Var | Fct | P-value | d.f. | Var | % Var | Fsc | P-value | d.f. | Var | % Var | Fct | |
| Australasia+CAP | 1 | 0.98 | 22.07 | 0.22 | 0.78 | 11 | 0.06 | 1.44 | 0.02 | 0 | 799 | 3.37 | 76.5 | 0.24 | 0 |
| HOK + WA + NZL | 2 | 0.32 | 8.57 | 0.09 | 0.03 | 10 | 0.01 | 0.4 | 0.004 | 0 | 799 | 3.37 | 91.03 | 0.09 | 0 |
| HOK + WA + NZL + CAP | 3 | 0.14 | 3.97 | 0.04 | 0 | 9 | 0.007 | 0.2 | 0.002 | 0.06 | 799 | 3.37 | 95.82 | 0.04 | 0 |
| HOK + Australasia + CAP | 2 | 0.09 | 2.48 | 0.02 | 0 | 9 | 0.007 | 0.2 | 0.002 | 0.06 | 786 | 3.39 | 97.32 | 0.03 | 0 |
| HOK+WA+NZL | 1 | 0.01 | 0.33 | 0 | 0.23 | 10 | 0.06 | 1.66 | 0.02 | 0 | 786 | 3.39 | 98.01 | 0.02 | 0 |

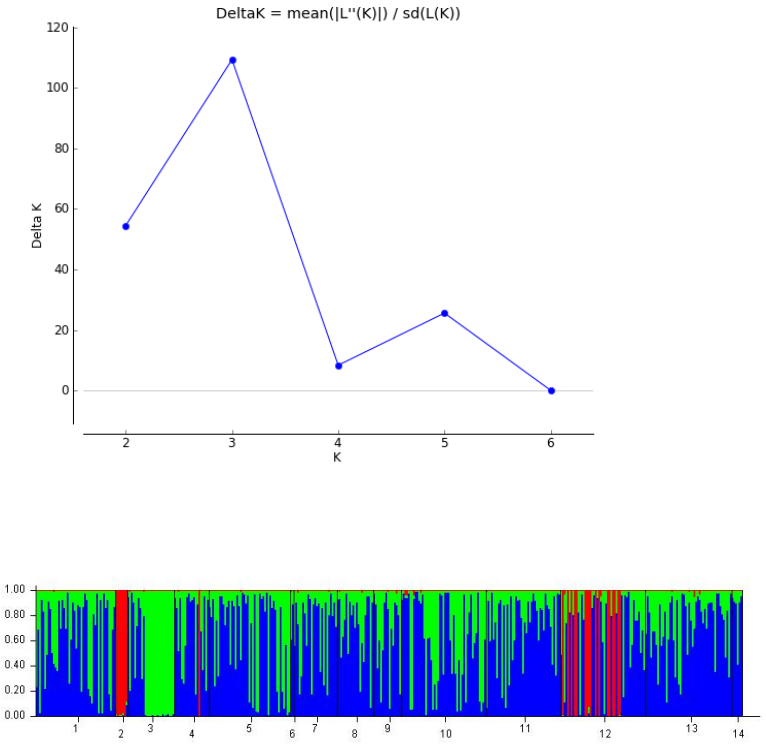



FIGURE 2.8: Results for the Bayesian structure analysis of *P. oxygeneios* assuming k=3 genetic clusters. Each line represents an individual. Individuals are ordered by site

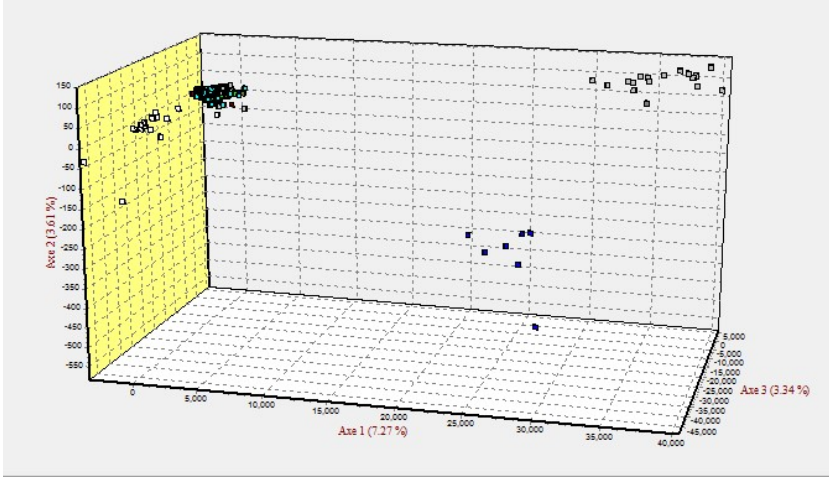



FIGURE 2.9: Factorial components analysis (FCA) of *P. oxygeneios* showing axes 1, 2, and 3. All grey individuals in the top right corner are from HOK. All blue individuals clustered together in the middle are from CAP. All other individuals cluster together in the top left corner.

### 2.4.3 Demographic history

#### Mitochondrial DNA data

High Harpending's Index and SSD values supported a model of sudden expansion, except in the case of CAP with P(SSD)=0.37 and P(R)=0.69 (Table 2.10). This was concordant with the results from the neutrality tests, with negative numbers produced by both Fu's F and Tajima's values. Only the Fu's F results were statistically significant, however Fu's F is considered a fairly reliable measure (Lane et al. 2016). The BSPs also supported the sudden expansion model. With the mutation rate of $2.0x10^{-8}$, the effective population size appeared to have increased substantially in the last 75,000 years, most prominently during the last 30,000 years. Using the mutation rate $3.6x10^{-8}$, population expansion appeared to have occurred 100,000-<25,000 years ago (ya), most prominently in the last 50,000 years. These results were in contrast to the mismatch distribution data, which deviated from the expected distribution using a model of population expansion. The multi-modal distribution observed in the figures was more reflective of an older, more stable population (Figure 2.10–2.12). Under a model of population expansion, a unimodal distribution was observed with the number of pairwise differences between haplotypes proportional to the frequency of occurrence. This occurred both when all samples were examined together, and when AUS and NZL were analysed separately. Excluding CAP, the timing since population expansion was estimated at 18,050-311,034 ya depending on the site and mutation rate. The general observation was that NZL expanded more recently (18,450-33,209 ya), with AUS having expanded earlier (130,3009-234,541 ya). HOK was similar to the rest of NZL (1,050-32,4 ya)(Table 2.10)

#### Microsatellite DNA data

The BOTTLENECK analysis provided no evidence for population contraction across all sites. All msatDNA loci conformed to the assumptions of mutation-drift equilibrium under all three models in the Wilcoxon sign-rank test. A normal L-shaped distribution was observed for all sites in the mode-shift test also, the expected pattern for a population in mutation-drift equilbirium (Appendix A).

TABLE 2.10: Estimation of population expansion using Fu's Fs, and mismatch distribution analyses: sum of squared deviation (SSD), Harpending's raggedness index,Tau ($\tau$) , population size before expansion ($\theta$o), population size after expansion ($\theta$1), and time since population expansion occurred (t) with P-values and confidence intervals in brackets.

| | Fu's Fs (P) | Tajima's D (P) | SSD (P) | Raggedness (P) | $\tau$ (95% CI) | $\theta$o (95% CI) | $\theta$1 (95% CI) | t |
|---|---|---|---|---|---|---|---|---|
| CAP | 1.653 | 0.41421 | 0.37 (0.00) | 0.69 (0.93) | Insufficient data | | | |
| WA | -24.72749 | -0.72981 | 0.02 (0.67) | 0.02 (0.76) | 88 | 1.225 (0-1.225) | 7.715 (2.349-19.395) | 234,5416-130,3009 |
| NZL | -19.43445 | -1.0228 | 0.018 (0.392) | 0.027 (0.368) | 1.246 | 6.121 (0-11.071) | 11.412 (6.242-99,999) | 33,209-18,450 |
| HOK | -5.968 | -1.02661 | 0.02 (0.32) | 0.05 (0.35) | 1.219 (0-10.87) | 2.52 (0-5.72) | 11.113 (4.64-99999) | 32,489-18,050 |
| ESP | -5.5 | -0.55794 | 0.02 (0.32) | 0.02 (0.62) | 11.67 (1.61-17.51) | 0.0 (0-4.58) | 10.80 (6.63-1496.60) | 311,034-172,797 |
| All sites | | | | | | | | |

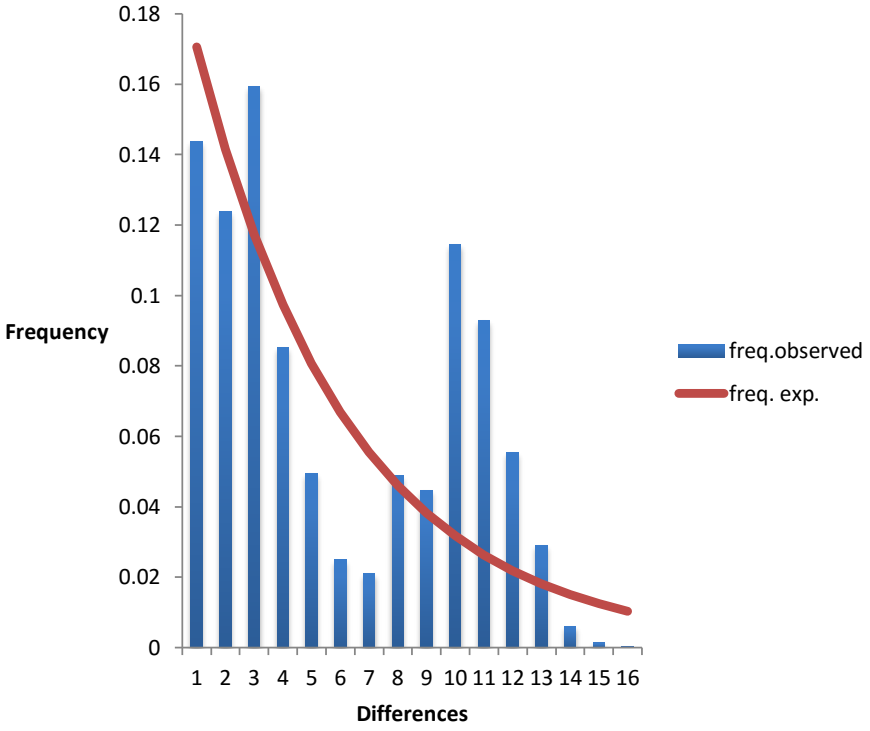



FIGURE 2.10: All sites, Mismatch distribution of all pairwise differences. Freq. expected is under a model of population expansion

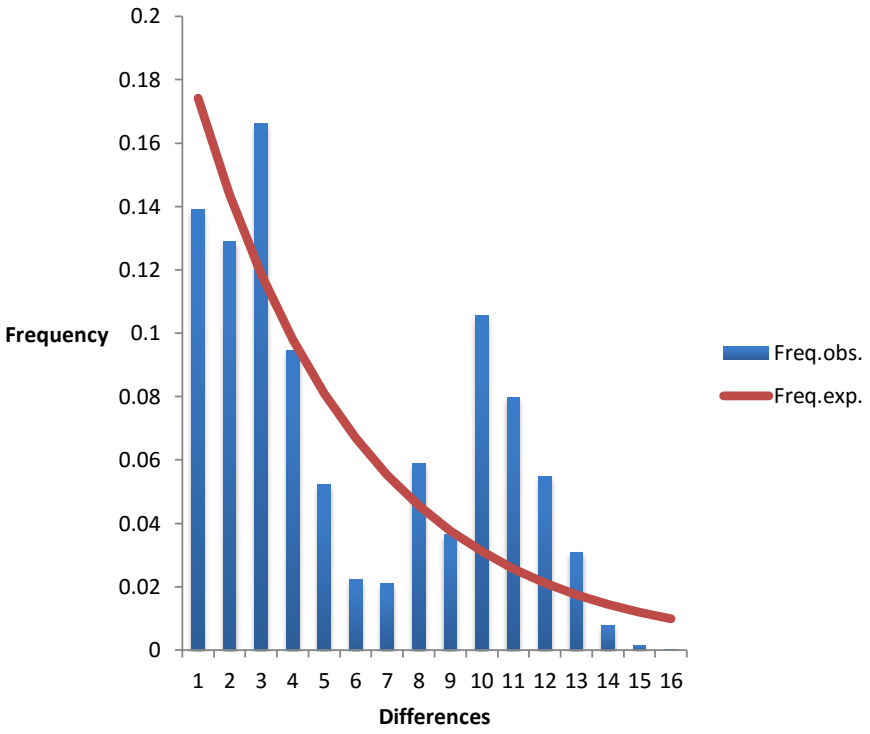



FIGURE 2.11: Mismatch distribution of all pairwise differences for NZL. Freq. expected is under a model of population expansion

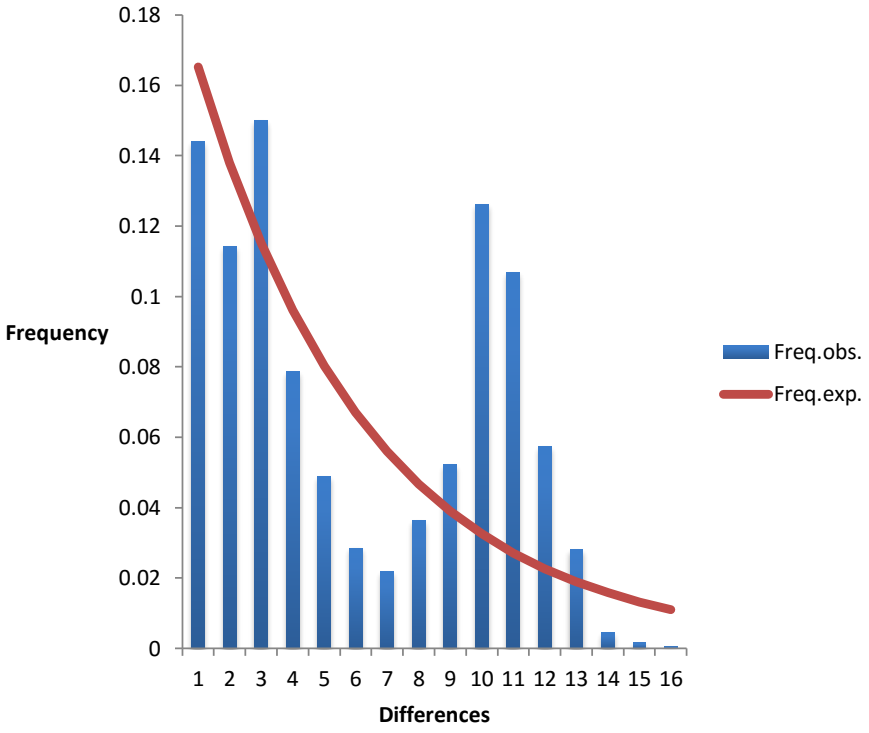



FIGURE 2.12: Mismatch distribution of all pairwise differences for AUS. Freq. expected is under a model of population expansion

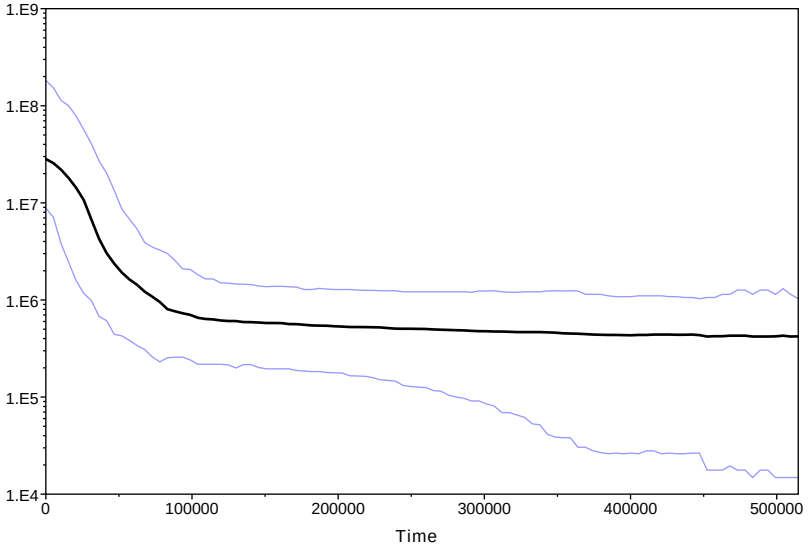



FIGURE 2.13: Bayesian skyline plot with $\mu$=2.0x10$^{-8}$. Time=time before present. Y axis=effective population size

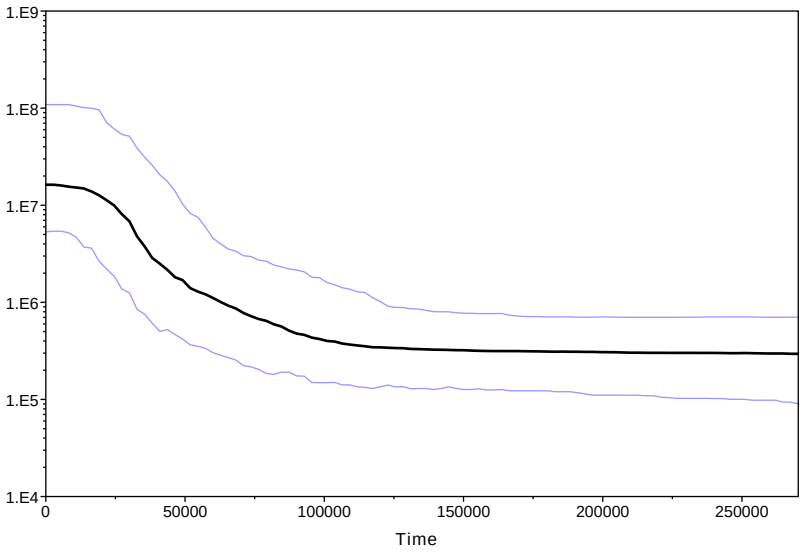



FIGURE 2.14: Bayesian skyline plot with $\mu$=3.6x10$^{-8}$. Time=time before present. Y axis=effective population size

## 2.5 Discussion

### 2.5.1 Genetic diversity and structure

The observed levels of genetic variation were high across all sites excluding CAP, with the average Hd (haplotype diversity) being 0.867. This result suggests that *P. oxygeneios* has a large historical effective population size, similar to those recorded for many other deep-sea fish including Orange Roughy, *Hoplostethus atlanticus* (Varela et al. 2012), Bluemouth, *Helicolenus dactylopteris* (Aboim 2005), Thornyhead Rockfishes, genus *Sebastolobus* (Stephien 2000), Cape Hake, *Merluccius capensis* (Von der Heyden 2007), Alaskan Pacific halibut, *Hippoglossus sterolepsis* (Nielsen 2010), and Alfonsino, *Beryx decadactylus* (Friess & Sedberry 2011). A large population size is a fairly typical feature of a marine species, which tend to have larger populations than their freshwater counterparts, partially due to their wider geographic ranges. Further, the environmental conditions tend to be less challenging over evolutionary time, which promotes gene flow and reduces the risk of population bottlenecks (DeWoody et al. 2000, McCusker & Bentzen 2010). Larger populations retain more genetic diversity because there is an increased effect in the number of mutations and a relative decrease in the strength of genetic drift (Shaw et al. 1999). The lower Hd (0.429) at Cape Town was most likely due to the small number of samples that were available from the site.

Concordance between results from both mtDNA and nuclear DNA is expected if a population has reached genetic equilibrium, and the historical barriers to gene flow reflect contemporary barriers to gene flow (Dammannagoda et al. 2011, Schloetterer et al. 2004). In the case of *P. oxygeneios* however, there was discordance between the AMOVA results from the mtDNA and msatDNA data. In the mtDNA analysis, between population variation was best described by Hokitika and the Chatham Islands representing separate populations to the rest of Australasia. Between population variation was best described when Hokitika and the Chatham Islands were grouped with the rest of Australasia in the msatDNA analysis. STRUCTURE, FCA, and some Fst results, supported Hokitika being grouped separately to the rest of Australasia. The difference between the Chatham Islands and the New Zealand sites in the mtDNA analysis was not found with any of the msatDNA analyses. One consistency among the analyses was the grouping of Cape

Town separately from all other sites, although the level of statistical significance was low.

To date, there have been no studies published that suggest the presence of more than one population of *P. oxygeneios*. Previous allozyme and tagging studies have failed to detect fine scale differentiation within New Zealand, however there has been speculation that Northland is potentially differentiated from the rest of New Zealand (Beentjes & Francis 1999, Smith & Johnson 1985). The results from the present study support the presence of a relatively panmictic population in Australasia. The extended juvenile phase (>4 years), highly mobile adults, and long life span makes pan-oceanic mixing of *P. oxygeneios* feasible. *P. americanus* also exhibit panmixia across ocean basins (Ball et al. 2000), and given the ecological similarities between the species, panmicitic structuring within Australasia is unsurprising. With Cape Town included in the Mantel test, there is a small correlation between genetic differentiation and geographic distance. However, once Cape Town is excluded from the analysis, any correlation disappears. It is likely that oceanic current systems play an important role in maintaining genetic connectivity between *Polyprion* stocks.

**Hokitika, Chatham, and the rest of NZL**

While the mtDNA data favoured a four population AMOVA model, with Chatham Islands and Hokitika grouped separately to the rest of Australasia, Australasia was kept as one all-inclusive group in the msatDNA AMOVA data. Under a model of neutrality and mutation-drift equilibrium, both markers should show similar results, even if it is not the same level of magnitude (DeBattista et al. 2012, Dammanagoda et al. 2011). One simple explanation for this discrepancy is that the differentiation observed in the mtDNA data represents historical differentiation and secondary contact between the previously isolated population has subsequently occurred, promoting admixture at the msatDNA loci. Another possible explanation is that a population bottleneck in the Hokitika area could have reduced heterozygosity at such a level that contemporary gene flow between Hokitika and other sites has yet to overcome the effects seen in the msatDNA FST data. These situations would not rule out contemporary demographic and genetic connectivity. It is possible that the effective population size of *P. oxygeneios* has been large enough to reduce the effects of genetic drift in the mtDNA relative to

the msatDNA. Dispersal of *P. oxygeneios* likely occurs for the most part as juveniles, reducing the possibility of a sex-ratio effect. However, these explanations do not adequately explain why Hokitika is grouped with the rest of Australasia in the AMOVA data, and grouped separately in the STRUCTURE and FCA analyses. While it could be argued that a lack of differentiation in the msatDNA reflects the ability of mtDNA to return to equilibrium more rapidly than msatDNA after a disturbance event, or the greater effects of homoplasy and gene flow on genetic signatures in msatDNA (DeBattista et al. 2012), these explanations are also contradicted by the discordance within the MsatDNA results. Of importance to note is the large portion of null alleles within the Hokitika site. Null alleles are common when msatDNA markers are developed in one species and used on another, because changes in the primer sites cause the amplification of some alleles to fail. This can exaggerate Fst and genetic distance measures. Although they may pose as a technical challenge, null alleles are relevant to population genetic studies as they represent an allele (Chapuis & Estoup 2007). Whatever the reason, there appears to be fine-scale differentiation between Hokitika and the rest of Australasia that is not detected in all analyses once areas are pooled for regional differences.

### South Africa

The data presented here, from both mtDNA and msatDNA markers, supported the differentiation of Cape Town from Western Australia and New Zealand, however the small sample size limits any firm conclusions. Robert (1986) suggested there was a third species of *Polyprion* based on a study using specimens collected off the coast of South Africa. He noted four juvenile *Polyprion* specimens with distinct spotted body colouration, and proposed that these four could represent a distinct species. Ball et al. (2000) also recorded distinctive mtDNA and msatDNA profiles from fish collected near South Africa, cited as unpublished data. It is difficult to determine whether the distinct DNA sequences and genotypes were new genetic variants that were part of a global population that was not properly sampled, samples from a genetically differentiated population, or a new species. The data presented in the current thesis research is insufficient for clarification of the species issue, however the results here do not reject the theory of a third *Polyprion* species by Ball et al. (2000).

**Demographics**

The largely panmictic structuring of *P. oxygeneios* suggested by this study (excluding Cape Town) reflects two potential scenarios 1) The sites in New Zealand and Australia represent one large, panmictic population with sites both genetically and demographically connected, and 2) Although there are enough migrants per generation to maintain genetic connectivity, some sites are demographically independent of each other. In this scenario, current genetic measures are inadequate at discerning this differentiation (Lowe et al. 2010). Nevertheless, the same traits that potentially promote the largely homogenous population structure observed could be the same traits that have resulted in small-scale differentiation observed within New Zealand. The dependence on oceanic currents for dispersal means that in cases where two sites are not connected by currents, gene flow is restricted. This results in differentiation patterns that are de-coupled with geographical distance, and has been observed in other species, for example, Atlantic cod, *Gadus morhua* (Ruzzante et al. 1998), Atlantic herring, *Clupea harengus* (Tatarenkov et al. 2007, Teacher et al. 2013), and cusk, *Brosme brosme* (Knutsen et al. 2009). Further, the K-selected biology of the species that gives individuals enough time to migrate with oceanic currents and promote gene flow, also leaves the species more susceptible to the effects of over-exploitation, a current concern for *P. oxygeneios*. (Wakefield et al. 2010) Over-exploitation can reduce genetic diversity and promote genetic drift.

The mismatch distribution, neutrality tests, and BSP's supported a sudden expansion model for Australasia. In this scenario, south-western Australia expanded earlier than New Zealand. This possibly suggests a migratory route from south-western Australia to New Zealand, however, given the uncertainty around the mutation rates used (3.6 $x10^{-8}$ and 2.0 $x10^{-8}$), this is purely speculative. The mismatch distribution suggested New Zealand (including Hokitika) expanded as recently as 18,000 ya, and up to 33,000 ya. In comparison, south-western Australia is an older population, having expanded between 130,000- 235,000 ya. The BSP's suggested population expansion in Australasia occurred sometime between the present day and 100,000 ya using the 2.0 $x10^{-8}$ mutation rate, and up until <25,000ya using the 3.6 $x10^{-8}$ mutation rate. A more species-specific mutation rate is needed to calibrate the time of population expansion to any historical events, however the estimated expansion times fall within the Pleistocene period (>10,000 years bp), which was punctuated by a series of large glacial-interglacial

changes, known to coincide with large bottleneck and expansion events in many coastal species. Deep-water species, such as *P. oxygeneios*, may be less affected by fluctuating environmental conditions associated with glacial cycles, however there is evidence that *P. oxygeneios* once inhabited shallower waters in New Zealand (Maxwell et al. 2011). In contrast, the haplotype and nucleotide data did not support the sudden expansion model. This was seen in the high number of mutational steps that were not represented by haplotypes, along with the lack of a star-like shaped network in the haplotype network. After losing genetic variation through a bottleneck event, a population will accumulate haplotype diversity faster than nucleotide diversity (Lane et al. 2016, Varela et al. 2012). When a population reaches mutation-drift balance, the number of haplotypes plateaus, and the ratio of haplotype diversity to nucleotide diversity decreases (Lane et al. 2016). For *P. oxygeneios*, both the nucleotide and haplotype diversity indices were high, indicating the presence of an older, more stable population. There was also no evidence of heterozygotic excess in Australasia, common in species that have undergone population contraction.

### Implications for fisheries management

Currently the New Zealand HPB (Hapuku-Bass) fishery is separated into eight stocks, which were divided up based on geographic features and concentrations of fishing effort. The stock boundaries themselves differ from the boundaries used to determine regional yields, or the tonnage of fish extracted from an area (Paul 2002). Cook Strait, for example, is situated on the borders of three stocks, HPB 2, HPB 7, and HPB8. While a maximum sustainable yield (MSY) is used to set the respective total allowable commercial catches (TACC) of each stock in order to reduce over-exploitation, it is currently unknown how relevant stock boundaries are to the natural population boundaries (MPI Science Group 2015). HPB stock boundaries might be over-split in New Zealand, with the results here suggesting, at most, there are three genetic populations in New Zealand. However, it is most likely that New Zealand is one panmictic population, with some level of differentiation occurring between Hokitika and other sites. While additional markers may reveal further levels of differentiation within the New Zealand EEZ, the data presented provides genetic support for the current management of Hokitika as a separate stock to the rest of New Zealand only.

*P. oxygeneios* is treated as a straddling stock in south-western Australia, targeted by both state and commonwealth fisheries. Annual commercial catches have increased since the first recorded state commercial landings in 1986, however the commonwealth-managed trawl fishery yields are significantly lower in comparison (Wakefield et al. 2010). Compared to the 2179 t caught New Zealand fishery (2002) the south-western Australian catch is small at 18 t (2010). The smaller Australian fisheries are partially due to nutrient poor waters that do not support high productivity (Paul 2002, Molony et al. 2012 ). Despite the low recordings and the risk status reported as moderate (status report), the current resource may already be fully exploited in south-western Australia (Wakefield et al. 2010). Further, there is an increasing shift in fishing effort from inshore to offshore fisheries (>100 m) after a decline in near shore stocks. This puts more pressure on species that were already sensitive to the effects of over-exploitation (Wise et al. 2007).The inclusion of samples from eastern Australia would be of particular value in future studies to determine whether any differentiation is observable between western and eastern Australia.

**Summary**

The Southern Hemisphere appears to consist of two discrete populations of *P. oxygeneios*, however without further sampling from South Africa this cannot be definitively stated. Although there is not enough data present to assess the taxonomic status of *P. oxygeneios*, the results from this study lend support to the suggestion in Ball et al. (2000) that there may be another species of *Polyprion* off South Africa. There is fine-scale differentiation within New Zealand, however the demographic implications of this differentiation are unknown. Future sampling from eastern Australia would be beneficial to the genetic analyses of *P. oxygeneios*. While catches of *P. oxygeneios* may be low compared to other species, the flesh is of high quality and in high demand (Paul 2002). Low landings in Western Australia may already reflect maximum yields, typical in K-selected species with lower reproductive output, slower growth rates, and longevity (Wakefield et al. 2010; 2013). Stocks in New Zealand appear to be over-split, which would have no biologically detrimental results.

# Chapter 3

# The phylogeographic structure and genetic differentiation of *Polyprion americanus* based on mitochondrial DNA control region sequences and microsatellite DNA genotyping

## 3.1 Introduction

While fishery managers may define a stock according to a specific area for pragmatic reasons (e.g. distribution of fishing effort), such definitions fail to incorporate information about the genetic and demographic substructuring of a population. By incorporating the substructure information when determining stock boundaries, managers are better able to offset the effects of fishery induced mortality with self-recruitment and immigration by introducing limits specific to individual fisheries and populations (Carvalho & Hauser 1995). While discrete stock structuring can easily be translated into fishery management boundaries, findings of no or little stock differentiation may be more difficult to interpret. Little differentiation could indicate a high level of movement and reproductive success, however we cannot be certain of this without knowing whether the population is in a genetic equilibrium (Chauhan & Rajiv 2010, Ward 2000).

Physical isolation among groups of individuals will lead to population structure when the evolutionary forces of genetic drift and adaptation begin to operate independently in each group (Cowen 2009). Traditionally, the wide distributions and lack of physical barriers that are typical for many marine fish species were thought to limit a population's ability for local adaptation because gene flow constantly homogenised variation (Hauser & Carvalho 2008). This view of marine species has changed as more examples of marine populations containing adaptive variation and relatively small effective population sizes suggests genetic drift might play an important role in marine species. For example, Atlantic Cod display varying maturation patterns linked to genetic variation at scales of tens of kilometres. This is despite gene flow of up to 10% migrants at each site (Olsen et al. 2008). Although levels of genetic differentiation are generally lower in scale in marine environments compared to species in the terrestrial environment, this differentiatiO has been shown to be demographically relevant (Hauser & Carvalho 2008).

There are a number of different methods that are used to estimate levels of migration, including non-genetic techniques (such as physical tags and chemical patterns from the environment), and genetic techniques (DNA sequencing or DNA microsatellites). However, genetic techniques are considered a reliable approach for estimating reproductive units when compared to non-genetic techniques. There are a range of genetic techniques, which can be distinguished from each other based on the mode of inheritance of the particular genomic region and the rate of

evolution (Hoffman et al. 2009). For example, mitochondrial DNA (mtDNA) does not undergo recombination because the paternal genome is excluded during fertilisation and only the maternal copy is inherited. This haploid mode of inheritance means all of the genes on the genome are linked and they have the same evolutionary history. In contrast, most nuclear loci evolve independently from each other because of recombination. (Hauser & Carvalho 2008, Schloetterer 2004).

The Giant sea bass *Polyprion americanus* is one of four species that belong to the family Polyprionidae. This species has an anti-tropical distribution that inhabits seamount, insular and continental slope waters (50-1000 m depth) throughout the Southern Hemisphere, but is limited to the Atlantic Ocean and adjacent water bodies in the Northern Hemisphere (Robert 1986). The juveniles have a long pelagic phase and only settle to become demersal adults at $>$ 50 cm total length. In the Southern Hemisphere, *Polyprion* stocks support fisheries off the Juan Fernandez Islands, Chile, New Zealand and Australia (Francis et al. 1999, Sedberry et al. 1999). Peres & Haimovici (2004) and Wakefield (2013) found the Southern *P.americanus* stocks to differ from previously recorded age and growth data from the North Atlantic, estimating a lifespan of $>$62 years (females) and $>$76 years (males). The differences between the northern and southern stocks almost certainly reflects differences in methodology. Ageing accuracy has improved, with nascent otolith preparation methods incorporating thinner sections (Wakefield, personal comms.). It is thought that a long pelagic life stage, coupled with an association with floating objects, facilitates pan-oceanic mixing of *P.americanus* populations (Robert 1986; Sedberry et al 1996; Machias et al 2003; Wakefield et al. 2010). In the Northern Atlantic, *P.americanus* form large spawning aggregations off of South Carolina, which would facilitate genetic panmixia. It is assumed that Southern Atlantic, Indian, and Pacific stocks also form spawning aggregations, however the locations of these aggregations are unknown (Paul 2002).

Using meristic and morphometric data, Robert (1986) concluded there were two species of *Polyprion* ; a deep bodied, uniform coloured *P. americanus*, and a shallow bodied, counter shaded *P. oxygeneios*. For the most part, Robert found adult *P. americanus* to be indistinguishable from each other, although he noted four *P. americanus* juveniles from South Africa with a distinct spotted body colouration. In contrast to the mottled and banded colouring in other *Polyprion* , this one character was interpreted as plesiomorphic. Based on the methodologies of the Evolutionary Species Concept, Roberts did not define these four South African

individuals as separate species. The potential for a third species of *Polyprion* has previously been noted by Ball et al. (2000) who reported distinctive mitochondrial DNA (mtDNA) sequences and microsatellite DNA genotypes in samples from South Africa, however their study did not include these samples in their genetic analysis due to a lack of accompanying data.

The aim of this study was, firstly, to use mtDNA and msatDNA markers to determine whether there is genetic differentiation within the Southern Hemisphere conistent with the findings of Ball et al. 2000, who proposed the presence of a third *Polyprion* species. Secondly, the demographic history of *P.americanus* was investigated. In general, mtDNA sequences provide insight into the demographic history of a species, while msatDNA markers are more suited for detecting contemporary differentiation (Hedgecock et al. 2007). The matrilinear inheritance pattern of mtDNA, along with its lack of recombination, reduces the effective population size to a quarter that of nuclear DNA. A smaller effective population size and the greater integrity of mtDNA relative to nuclear DNA are both characteristics that make mtDNA more sensitive to demographic changes, and useful for studying differentiation over longer evolutionary times. Compared to mtDNA genes, msatDNA loci are typically independent loci because of recombination. They also have a high mutation rate, making them ideal markers for estimating levels of population genetic variation and differentiation.

## 3.2 Methods

### 3.2.1 Sample collection

Tissue samples of *Polyprion americanus* were collected by researchers from recreational and commercial catches in three areas in the Southern Hemisphere; Argentina, south-western Australia (i.e. Perth, South West Capes, Albany, Esperance, Table 3.1), and south-eastern Australia (Barcoo and Taupo Banks). Date, location and latitude/longitude data, collector (recreational or commercial), total length (mm, TL), depth (m), and sex were recorded for most of the individuals, with those samples of unknown location and/or latitude/longitude information excluded from sequencing (Table 3.1). Samples were first stored in a 5ml vial containing absolute ethanol (99.9%); the ethanol was then drained and replaced

by dimethyl sulphoxide (DMSO) for shipping to New Zealand (NZ), because it is less hazardous. Upon arrival in NZ, all samples were promptly drained of DMSO, and returned to ethanol (80%). Samples were stored at 4°C until needed.

It has been suggested that a sample size of 50 individuals will enable sampling of 99% of the genetic variation within a population (Smith 2008). However, these sample sizes were difficult to obtain from South America due to a prohibition on the retention of *Polyprion*. To limit sampling bias, an attempt was made to balance the sample sizes from each location. This meant that despite more samples being available, fewer than 50 individuals from Perth Canyon were DNA sequenced (Table 3.1). The data collected from South America, western and eastern Australia were compared to mtDNA primers previously reported in Lane (2013).

TABLE 3.1: *Polyprion americanus* samples

| Sample location | Sample size (n) | Sample size included in final analyses |
|---|---|---|
| Albany (WA) | 2 | 2 |
| Esperance (WA) | 23 | 23 |
| Perth Canyon (WA) | 61 | 46 |
| SW Capes (WA) | 54 | 47 |
| Barcoo bank (EA) | 18 | 17 |
| Taupo bank (EA) | 35 | 25 |
| South America | 22 | 20 |
| New Zealand | 25 | 25 |
| Total | 215 | 205 |

WA=Western Australia, EA=eastern Australia.

### 3.2.2 Mitochondrial DNA extraction, amplification and sequencing

Mitochondrial DNA was first extracted using a standard phenol-chloroform protocol (Sambrook et al. 1989) and PCRs were then performed in a TGradient thermal Block (Biometra, Goettingen, Germany). All reaction mixes consisted of a 15 $\mu$l mix of 67 mM Tris-HCL pH 8.8, 16 mM $(NH_4)_2SO_4$, 3.0 mM $MgCl_2$, 200 $\mu$M of each dNTP, 0.5 $\mu$M of each primer, 0.6 $\mu$g/$\mu$L Bovine Serum Albumin (BSA), and 1 unit of BIOTAQ$^{TM}$ DNA Polymerase (Bioline). Mitochondrial DNA control region primers specifically designed previously for *Polyprion* were the forward primer tRNA-Pro-HPB (5'-CCTACCCCTAACTC-CCAAAGC) and reverse primer CCD-BASS (5'-AAGAGAACCCCTTGCTCGCTG), designed specifically

for *P.americanus*. The amount of template DNA used varied from 0.5$\mu$l- 1.5$\mu$l. Thermal cycling consisted of 36 cycles: 95°C for 2 minutes (mins), followed by 36 cycles of 94°C for 30 seconds (s), 61°C for 30 s and 72°C for 45 s and a final extension at 72°C for 10 mins. The resultant 550 base pair (bp) amplicons were electrophoresed in 1.0% agarose gel, stained with ethidium bromide, and visualised under a UV-light source. Amplicons were purified using Exo-SAP-IT as per the manufacturer's instructions before being sent to either Massey University Genome Services, New Zealand, and Macrogen, Seoul, Korea, where DNA sequencing was performed using the forward primer on an ABI 3730 Genetic Analyser.

### 3.2.3 Microsatellite amplification and genotyping

Four 96-well plates containing 384 specimens were prepared and sent to AgResearch for genotyping. Of the 384 specimens, 194 were *P. americanus*. To ensure consistency, the samples used were from the same individuals as those for the mtDNA sequencing, and an attempt was made to have samples from each site and species equally represented in the data. However in some cases, such as with the site Argentina (ARG), this was impossible due to a low number of samples. *Polyprion* samples that were successfully extracted and sequenced were preferred over samples from individuals that did not sequence as well.

Small amounts of tissue, no greater than 1.0 cm long and 0.5 cm wide, were removed from the *Polyprion* tissue samples and placed in wells with enough ethanol (80%) to fully cover the tissue pieces. Once each 96-well plate was filled and lids tightly sealed, samples were kept at 4°C until shipping to AgResearch (GenomNZ) in Invermay, New Zealand. Genomic DNA was extracted using a Chelex® procedure (Bio-Rad laboratories) and a combination of five microsatellite DNA loci, previously published by Ball et al. (2000) and four propriety loci previously developed by AgResearch for NIWA (table 3.2) and use in *P.oxygeneios* were used in a multiplex genotyping panel. The multiplex was conducted with an annealing temperature of 56°C and a MgCl2 concentration of 20.nM. The primer concentration ranged from 0.1 to 0.6 $\mu$M. One of the primers in each pair was fluorescently labelled and an ABI3730 Genetic Analyser (Applied Biosystems) was used to determine the allele sizes.

TABLE 3.2: Nine microsatellite DNA markers and type of repeat

| Marker | Nucleotide Repeat |
|---|---|
| Pam010 | Di (GT) |
| Pam017 | Di (GT) |
| Pam021 | Di (AC) |
| Pam025 | Di (CA) |
| Pam035 | Di (GT) |
| GJLKPX | Di (AC) |
| GJSLB2 | Di (AG) |
| GGOQ6A | Di (TC) |
| GH0OIK | Tri (TTG) |

## 3.3 Genetic analyses

### 3.3.1 Genetic diversity

#### Mitochondrial DNA data

The final data set consisted of 205 *P. americanus* samples. DNA sequences were edited and aligned using GENEIOUS 6.1 (Biomatters). Summary statistics, including the number of segregating sites (S), haplotypes (H), nucleotide diversity, haplotype diversity (Hd), private haplotypes (Hp) and the average number of pairwise differences (K), were collected using DNASP 5 (Librado & Rozas 2009). Due to the small sample size collected from Albany (n = 2), it was left out of subsequent analyses. Fu's F statistics, Tajima's D, and pairwise fixation index ($\Phi_{ST}$), were measured in ARLEQUIN 3.5 (Excoffier et al. 2005).

To visualise the relationship between haplotypes, along with the frequency of each haplotype, Minimum Spanning Networks were generated in POPART (www.popart.otago.ac.nz) using 1000 iterations. Haplotypes with a frequency <0.04 were considered rare. Rarefaction curves were constructed for each site to check whether the diversity observed accurately reflected the diversity of the wider population. ANALYTIC RAREFACTION 1.3 software (Holland 2003) simulates the number of haplotypes expected with each increment in sample size. The point at which a plateau is reached is meant to indicate the point at which a sample size contains a true representation of population diversity.

**Microsatellite DNA data**

Evidence for null alleles and score errors was assessed using MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) with 1,000 iterations. Deviations from the Hardy Weinberg equilibrium (HWE) were tested for across all sites and loci using the exact test with a $10^6$ Markov chain length and $10^5$ de-memorization steps implemented in ARLEQUIN 3.5 (Excoffier et al. 2005). A pairwise test (likelihood ratio) of linkage disequilibrium (LD) was also carried out in ARLEQUIN 3.5 with 10,000 permutations. For both tests, significance values were estimated at 0.995 confidence after Sequential Bonferroni corrections.

Basic diversity indices (TABLE 3.3) calculated using Genepop included observed and expected heterozygosity (Ho and He), number of alleles per locus, and the inbreeding coefficient (Fis). Allelic richness, private allelic richness and number of private alleles were calculated using Hp-rare. Allele frequency distributions were then generated for each population and locus using GENALEX 6.5 (Peakall & Smouse 2006: Peakall & Smouse 2012).

An FST outlier detection method was performed in LOSITAN to test the theory of microsatellite loci neutrality (Antao et al. 2008). This method is able to identify candidates under selection by analysing the relationship between FST and heterozygosity (Beaumont & Nichols, 1996). Theoretically, loci under selection, or loci that are linked to genes under selection, will appear as outliers when compared with those influenced by drift and gene flow only. Both a stepwise (SMM) and infinite allele (IAM) mutation model were performed using 100,000 simulations and a confidence interval of 0.995. Then, to ascertain whether sample sizes were large enough to contain a level of diversity representative of the population, allele rarefaction curve graphs were generated. First, Hp-rare was used to generate a number of discovered alleles for cumulative sampling size with each increment being one individual. Secondly, the relationship between 'discovered' alleles and cumulative sample sizes was investigated.

### 3.3.2 Population structure

#### Mitochondrial DNA data

Analysis of Molecular Variance (AMOVA) was used to estimate the level of genetic differentiation within and among sampling locations. Using ARLEQUIN 3.5, the following four population groupings of *P. americanus* were tested, each with 10,000 permutations: The first simulation consisted of four 'populations'; 1) Argentina (ARG), 2) Perth, South West Capes, Albany, and Esperance, (PER, SWC, ALB, ESP), referred to herein as 'WA', 3) Barcoo and Taupo Banks (BAR, TAU), or 'EA', and 4) New Zealand (NZL). It was suspected that one or two of the 'populations' may have accounted for most of the 'between group' variation, so another three groupings were tested, the first with three 'populations': 1) WA, 2) EA, and 3) NZL. The other two consisted of two 'populations' only, 1) NZL and all Australian sites (AUS), and 2) ARG and all AUS sites. All groupings were based on potential genetic breaks identified elsewhere, and large geographic gaps.

Pairwise genetic differences ($\Phi_{ST}$) were used to estimate the level of gene flow between all sites using 1000 permutations in ARLEQUIN 3.5. To determine whether a pattern of Isolation-By-Distance was present, linearised ΦST values were plotted against geographic distance (km) using GOOGLE EARTH 7.1. The distances were based on the shortest possible distance, by sea, between two sites. Due to the small sample size of TAU, this was again excluded from analyses. Coordinate measurements were based on a point equidistant to the two sample locations. Statistical significance of the linear regression was assessed using a Mantel test conducted in ARLEQUIN 3.5 using 10,000 permutations. Bonferroni corrections were made to reduce the chances of a type I error (incorrectly rejecting our null hypothesis).

#### Microsatellite DNA data

Three approaches were used to measure genetic distance and estimate levels of genetic differentiation between sites. First, Weir & Cockerham's (1984) Fst analysis was implemented using FSTAT 2.9.3.2. (Goudet 1995). Fst estimates were used without assuming Hardy Weinberg equilibrium, 1000 iterations, 95% confidence intervals, and Bonferroni corrections. Then, both G'st (Hedrick 2005) and Jost's D estimator, Dst (Jost 2008), were estimated using GENODIVE (Meirmans &

Van Tienderen 2004). All three measures were used to account for the limitations of each measure on their own. Fst, while widely used, bases genetic distance on decreased heterozygosity, which, in multi-allelic data, could theoretically be high even if sites contained a completely different set of alleles. With highly polymorphic microsatellite data, this can result in a downward bias in Fst values. While both G'st and Jost's D account for variation in the maximum obtainable distance value, they are both often biased upwards instead. Jost's D is measured using the effective number of alleles rather than heterozygosity, however, it is also sensitive to changes in mutation rate. Due to a lack of data at loci PAM010, PAM025, and GH0OIK in ARG, these three loci were excluded from distance measures when ARG was included in analyses. Linearised Fst values were later plotted again the shortest geographical distance (km) by sea to test for a pattern of Isolation-by-distance (IBD) in all samples, excluding ALB. ARLEQUIN 3.5.1 was also used to determine significance using a Mantel test with 1,000 permutations.

Analysis of Molecular Variance (AMOVA) was then implemented in ARLEQUIN 3.5 using the following groupings: 1) ARG and Australasia grouped separately; 2) ARG, AUS, and NZL grouped separately; 3) ARG, WA, EA, and NZL grouped separately; and 4) ARG, WA, and the rest (EA and NZL) grouped separately.

Using STRUCTURE 2.3.4 (Pritchard et al. 2000), a Bayesian Clustering Analysis was performed. Each cluster (K), which was set to a range of 1-7 and simulated 10 times, is characterised by a set of allele frequencies which each individual is then assigned to. Markov chain Monte Carlo (MCMC) simulations were run 10,000, 100,000, 400,000, and 1,000,000 times, after a burn-in period of $10^5$ steps, under both admixture and non-admixture models with the 'locprior' function implemented. Delta K, used to estimate the value of K that best reflects the population's structure, was measured using STRUCTURE HARVESTER 0.6.94 (Earl & VonHoldt 2012). A Factorial Component Analysis (FCA) was also conducted in GENETIX (Belkhir et al. 2000) to visually examine the differences between each individual in the sample using three main axes.

### 3.3.3 Demographic history

#### Mitochondrial DNA data

The demographic history was investigated using three approaches. Firstly, the neutrality tests Fu'f F (Fu 1997) and Tajima's D (Tajima 1989) were calculated in ARLEQUIN 3.5 using 1000 simulations to measure the number of segregating sites in the data against the observed nucleotide diversity. Significant negative values ($P<0.02$) indicate an excess of segregating sites given the observed nucleotide diversity, expected in a population that has recently undergone expansion (Fu 1997, Tajima 1989).

The second approach was Mismatch distribution analysis, conducted in DNASP 5. Here, the frequency of pair-wise differences between haplotypes is measured. The demographic history of a population may be inferred by the shape of the distribution, which is expected to change under different scenarios. For example, a unimodal distribution is expected in a population that has recently undergone population expansion, whereas a multimodal or erratic distribution reflects more long term stability (Rogers & Harpending 1992). To fit the data to a sudden expansion model (1000 replicates), both Harpending's raggedness index (Harpending 1994) and the sum of squared deviations (SSD) were calculated using ARLEQUIN 3.5. The calculations were made by both treating all sites as one population for each spp., as well as separating ARG, WA, EA, and NZL. To examine the changes in population size, the demographic parameters $\tau$, and $\theta$ were estimated from the mismatch distribution in ARLEQUIN 3.5, and the equation t= t/2u (Rogers & Harpending 1992), where $\mu$ is the mutation rate of the sequence (not per nucleotide), was used to estimate a time since population expansion. The value of $\mu$ can be determined using the formula $\mu = 2\mu$, where $\mu$ is the mutation rate per nucleotide site and k is the number of analysed nucleotide base pairs. The mutation rates of $2.0\text{x}10^{-8}$ proposed by Brown et al. (1979) for the vertebrate mitochondrial genome based on mammalian data and $3.6\text{x}10^{-8}$ calibrated by Donaldson & Wilson Jr. (1999) for snook (Percoidei; Centropomidae) were used, both because no mutation rate has been calibrated for *Polyprion*, and to remain consistent with Lane et al. (2016), who used both mutation rates also.

Lastly, Bayesian skyline plots were constructed in BEAST 1.8.3 (Drummond et al. 2012) for Australasia, excluding Argentina and South West Capes due to the

differentiation observed between these sites and others. This method quantifies the relationship between the genealogy of the sequences and the demographic history of the population (Ho & Shapiro 2011). The Marko Chain Monte Carlo simulations were run for $10^7$ iterations with the HKY substitution model, and a strict molecular clock. Two mutation rates were used; $2.0\text{x}10^{-8}$ and $3.6\text{x}10^{-8}$. The results were checked, and skyline plots were constructed in TRACER (Rambaut et al. 2014).

### 3.3.4 Mutation-drift equilibrium and bottleneck

#### Microsatellite DNA data

To determine whether samples were in mutation-drift equilibrium, a mode test was implemented using BOTTLENECK 1.2.02 (Piry et al. 1999). This method compares the proportion of alleles at low frequencies to an expected L-shaped distribution in a population at equilibrium. When population sizes decrease and the strength of genetic drift is stronger, alleles at low frequencies are eliminated from the population sooner than those at a higher frequency, skewing the distribution. Allele frequency distribution plots were created using GENALEX (Peakall & Smouse 2006, 2012). BOTTLENECK was again used to run a Wilcoxon sign-rank test, which is also based on measures of heterozygosity. Three mutation models were used for the simulation: (i) the Stepwise Mutation Model (SMM), ii) the Infinite Alleles Model (IAM), and iii) the Two Phase Model (TPM). Analyses were performed with 1000 iterations, and the TPM was set to 95 % single-step mutations and 5% multi-step mutations.

## 3.4 Results

### 3.4.1 Genetic diversity

#### Mitochondrial DNA data

A total of 205 *P. americanus* DNA sequences from the mtDNA control region were analysed in this study (TABLE 3.3). Although samples were taken from

eight different sample sites in South America, Australia and New Zealand, more than half of the samples (n=118) were from south-western Australia, and a large number (n=42), from eastern Australia. Just 45 of the samples were from South America and New Zealand. The sequences were aligned, and trimmed to 495bp long. The sequences showed asymmetric base frequencies of A=0.34, C=0.21, G=0.14, T=0.31. In total, there were 56 Polymorphic sites, with 45 of these being parsimony informative. There were 54 polymorphic sites and two indels, and the other two contained three indels. Hd ranged from 0.83824 in Barcoo Bank to 0.95652 in Esperance. In total, 67 haplotypes were present, 19 of which are present in samples from the Perth Canyon. No haplotype was found at every site, however the most abundant was haplotype 2, which was recorded at all sites except ARG. All the mtDNA sequences found at the ARG site were private haplotypes (14).

TABLE 3.3: Summary statistics for *P.americanus*

| Location | | long | lat | N | h | Hd | S | K |
|---|---|---|---|---|---|---|---|---|
| South West Capes | SWC | 114.465 | -33.6931 | 47 | 18 | 0.84921 | 24 | 4.18316 |
| Albany | ALB | 118.8082 | –35.0415 | 2 | 2 | 1 | 1 | 1 |
| Perth | PER | 115.134 | –31.9743 | 46 | 19 | 0.88889 | 28 | 4.67536 |
| Esperance | ESP | 122.1227 | –34.5 | 23 | 16 | 0.95652 | 22 | 5.73913 |
| Barcoo | BAR | 156.25 | –32.58333333 | 17 | 10 | 0.83824 | 24 | 4.64706 |
| Taupo | TAU | 156.1666667 | –33.16666667 | 25 | 15 | 0.93333 | 25 | 4.85333 |
| New Zealand | NZL | 173.536668 | –34.989566 | 25 | 14 | 0.92333 | 26 | 5.40667 |
| Argentina | ARG | –62.666 | -47.695 | 20 | 14 | 0.93684 | 17 | 4.35263 |

N=number of sequences, h=number of haplotypes, Hd=haplotype diversity, S= number of segregating sites, K=number of pairwise differences, Hp=number of private haplotypes, and $\pi$=nucleotide diversity

The rarefaction analysis did not reach a plateau at any site, indicating that more haplotypes would be discovered with a greater sampling size. The haplotype combination and haplotype frequencies for each site (excluding ALB) are shown in Figure 3.3. Note that the colours used in each pie are consistent with the colours used in the Minimum Spanning Network. While haplotypes with a frequency $<0.04$ are considered rare and are generally grouped together, in the case of ARG, every private and rare allele is represented by a different shade of grey so that the number of haplotypes present can still be seen.

The Minimum Spanning Network revealed the presence of one main haplotype, i.e. haplotype 2. However, haplotype 16 was almost as abundant, and only found in

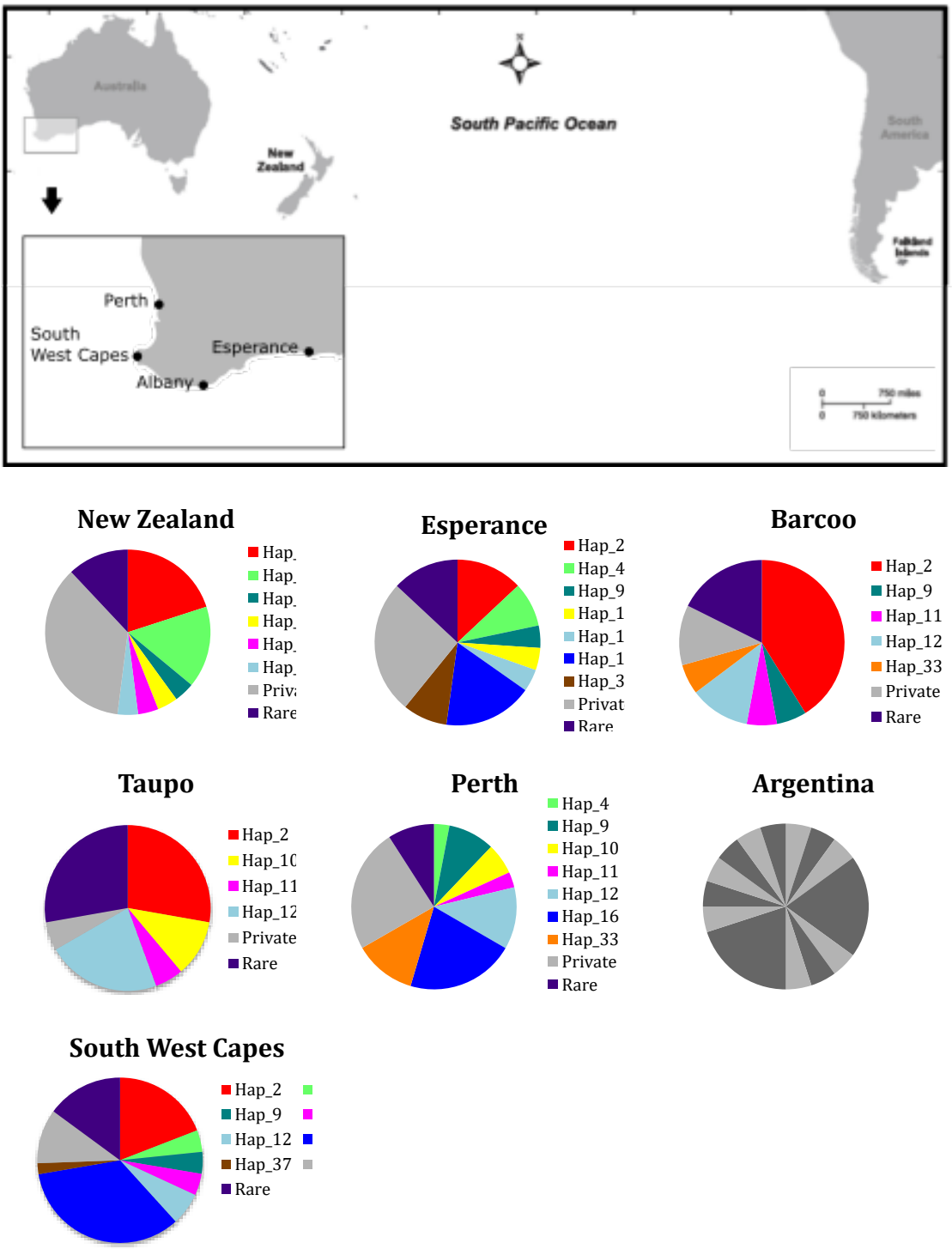



FIGURE 3.1: Sampling locations and haplotype distribution for *P.americanus*

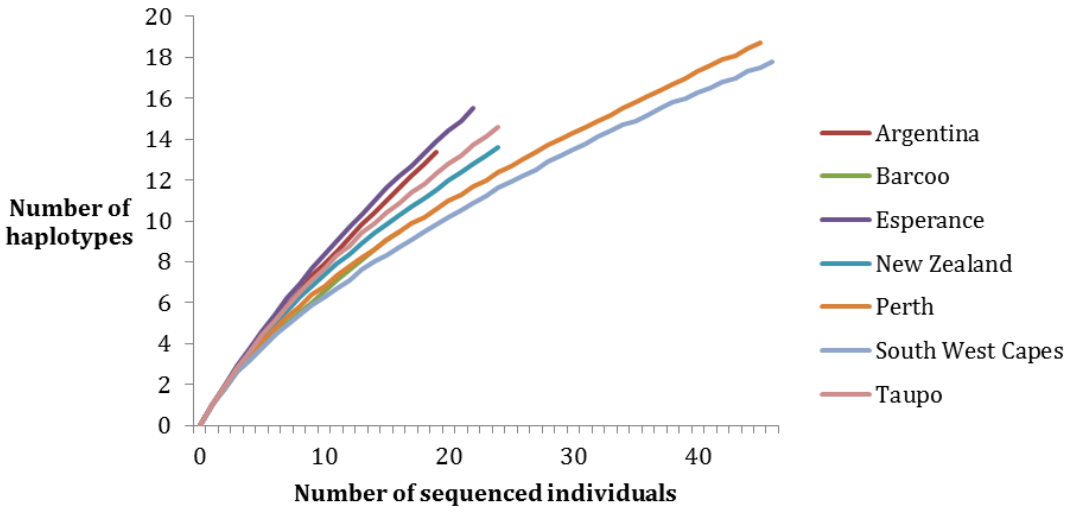



FIGURE 3.2: Haplotype discovery curve.

samples from south-western Australia. Haplotype 12 was the next most abundant haplotype which, like haplotype 2, is seen at all sites excluding ARG. The clustering of private alleles in the network were completed segregated between ARG and those from all other locations. The closed circles in the network correspond to missing (non-sampled) mutational steps.

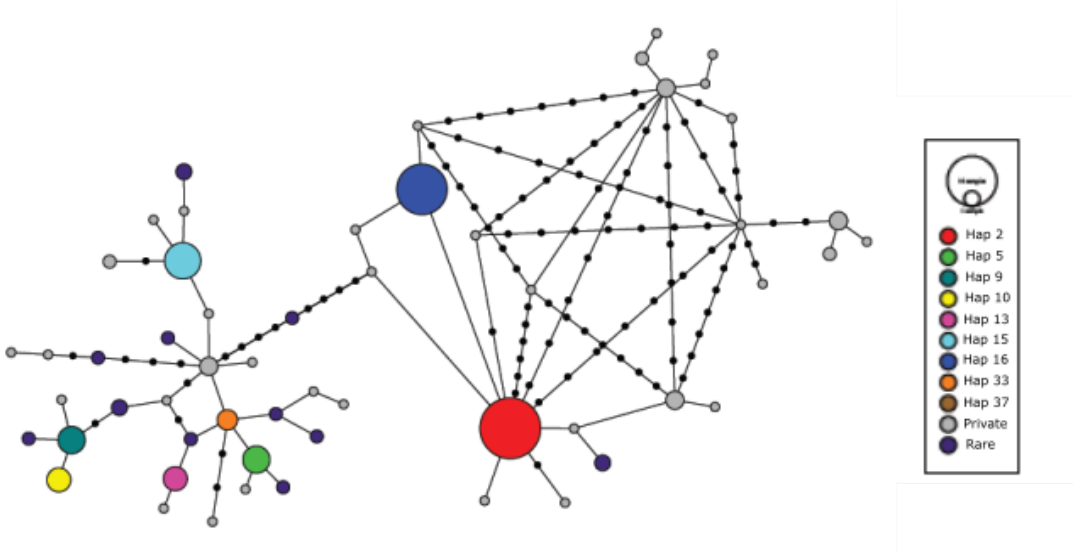



FIGURE 3.3: Haplotype genealogy of the mtDNA control region sequences. Each colour represents a unique haplotype. The scale shown indicates the frequency of each haplotype. Small, black circles represent putative mutational steps between haplotypes

### Microsatellite DNA data

Of the 198 *P. americanus* samples that were genotyped, 176 were successfully genotyped at eight or nine microsatellite DNA loci, and 22 samples were successfully genotyped at six loci. Of the 198 samples, two were from tissue samples collected in ALB, 18 from BAR, 23 from ESP, 61 from PER, 51 from SWC, 34 from TAU, 22 from ARG, and just nine from NZL. Individuals were removed from the data set when data was missing at three or more loci, except those from Argentina (ARG). Samples from ARG were analysed in microchecker using only six of the loci due to missing data in most ARG samples at sites PAM010, PAM025, and GH0OIK, which possibly reflects actual genetic differences, such as a mutation at the primer site, rather than genotyping errors. Testing for homozygote excess at 95% confidence provided no evidence of null alleles, except at one locus, at one site (PER He=29.88, Ho=34). Having just two samples from ALB provided an insufficient level of data to perform analyses for this site, and so was excluded from HW and LD analyses, as well as rarefaction analyses. Summary statistics, including observed and expected heterozygosity (Ho and He), allelic richness (Ar), fixation index (Fis), and number of private alleles (PA) are provided in Table 3.4.

Table 3.4: Summary statistics for nine microsatellite loci.

| Sample site | N | H0 | He | Ar | PAr | Fis | PA |
|---|---|---|---|---|---|---|---|
| Argentina | ARG | 19 | 4.111 | 4.332 | 3.15 | 1.04 | 0.021 |
| Albany | ALB | 2 | NA | NA | NA | NA | NA |
| Esperance | ESP | 19 | 7.556 | 7.047 | 1.9 | 0.1 | -0.065 |
| Perth | PER | 46 | 14.625 | 16.313 | 1.88 | 0.08 | 0.087 |
| South West Capes | SWC | 47 | 16.556 | 17.139 | 1.86 | 0.09 | 0.028 |
| Barcoo banks | BAR | 18 | 7.444 | 7.751 | 1.93 | 0.14 | 0.004 |
| Taupo banks | TAU | 35 | 12.444 | 11.829 | 1.8 | 0.08 | –0.072 |
| New Zealand | NZL | 12 | 4.444 | 4.598 | 1.99 | 0.15 | 0.019 |

N=sample size (N), H0=observed heterozygosity, He=expected heterozygosity, Ar=allelic richness, Fis=fixation index, PA= number of private alleles

Sample sizes in this study were probably insufficient to capture the true level of *P. americanus* diversity in the Southern Hemisphere. The allele rarefaction curves for sites with the smallest sample sizes (ARG and NZL) do not reach plateau at many of the loci, as can be seen at the most polymorphic locus, GJLKPX (see Appendix B). At this locus, none of the curves reach a true plateau, however, as can be seen in other loci, sites with the greatest sampling (SWC, TAU, and

PER) appear to be almost at plateau. All rarefaction curves can be found in the appendix.

After applying Bonferroni correction for multiple tests (p=0.0093, p(ARG)=0.0014), all loci were found to be in Hardy-Weinberg equilibrium. Likewise, there was no significant LD between any loci after correction for multiple tests (P=0.0014, P(ARG)=0.0033). However, locus GJLSB2 was a candidate for positive or balancing selection under both a Stepwise Mutation (SMM) and Infinite Alleles Model (IAM). PAM035 also showed signs of selection under SMM, but not under an IAM (Figures 3.6 & 3.7). At these two loci the assumption of neutrality is rejected, thus results from these two loci should be analysed with caution.

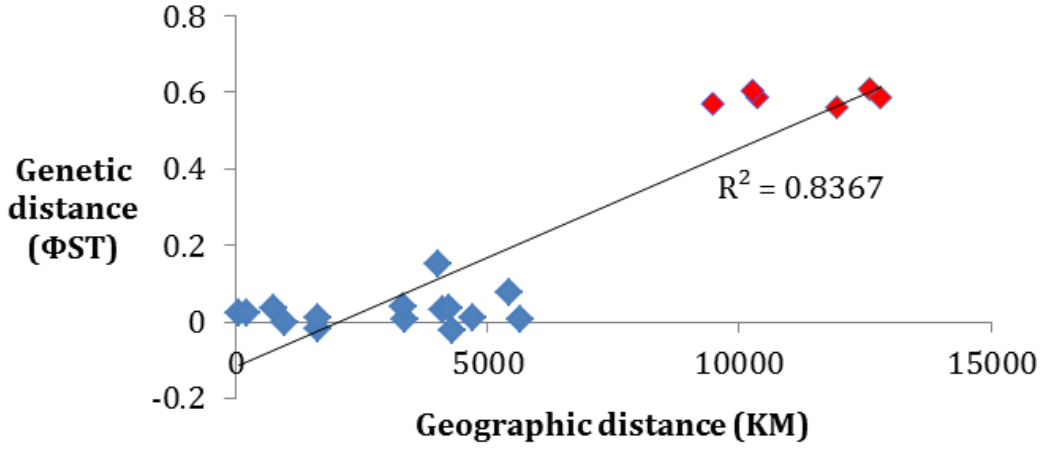



FIGURE 3.4: Correlation between $\Phi_{ST}$ and geographic distance

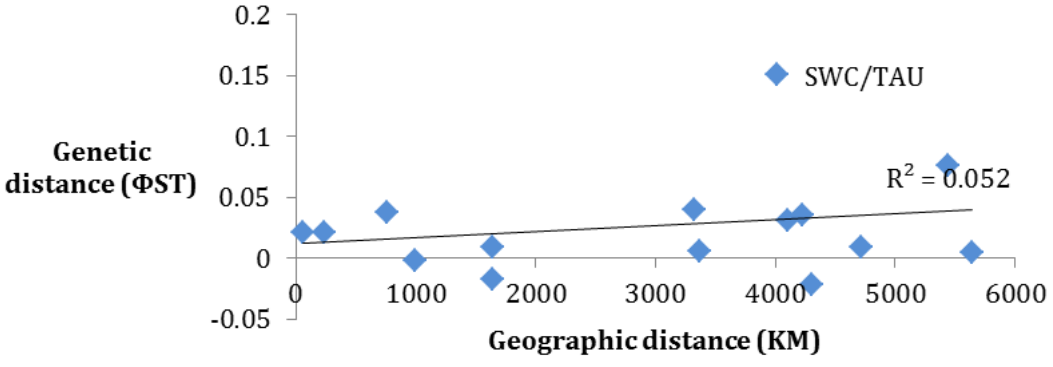



FIGURE 3.5: Correlation between $\Phi_{ST}$ and geographic distance, excluding ARG

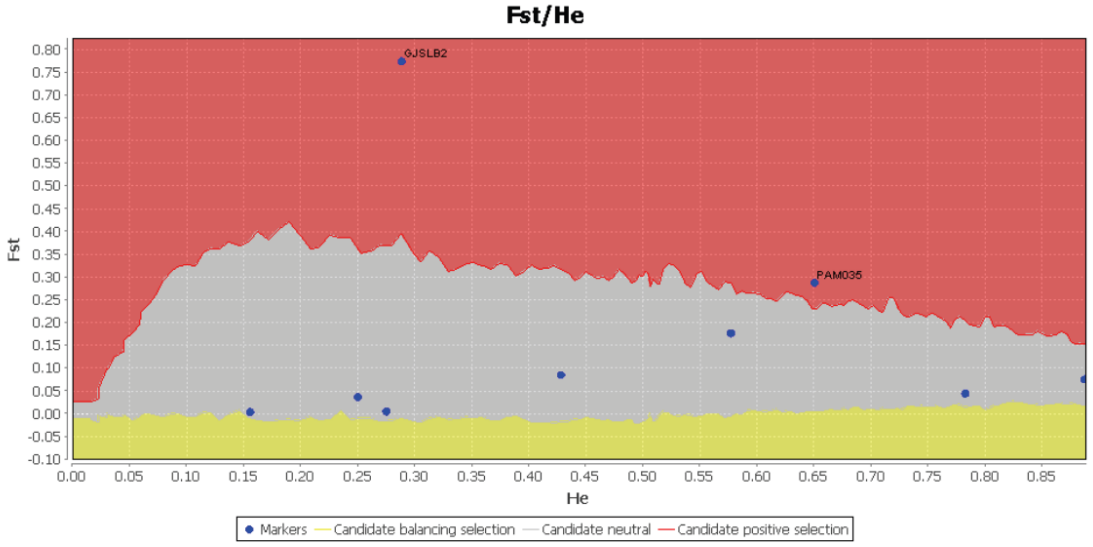



FIGURE 3.6: Fst outlier results under the Stepwise Mutation Model (SMM). Both GJLSB2 and PAM035 are candidates for positive selection with He=0.289 and 0.650 respectively, and Fst=0.774 and 0.288 respectively

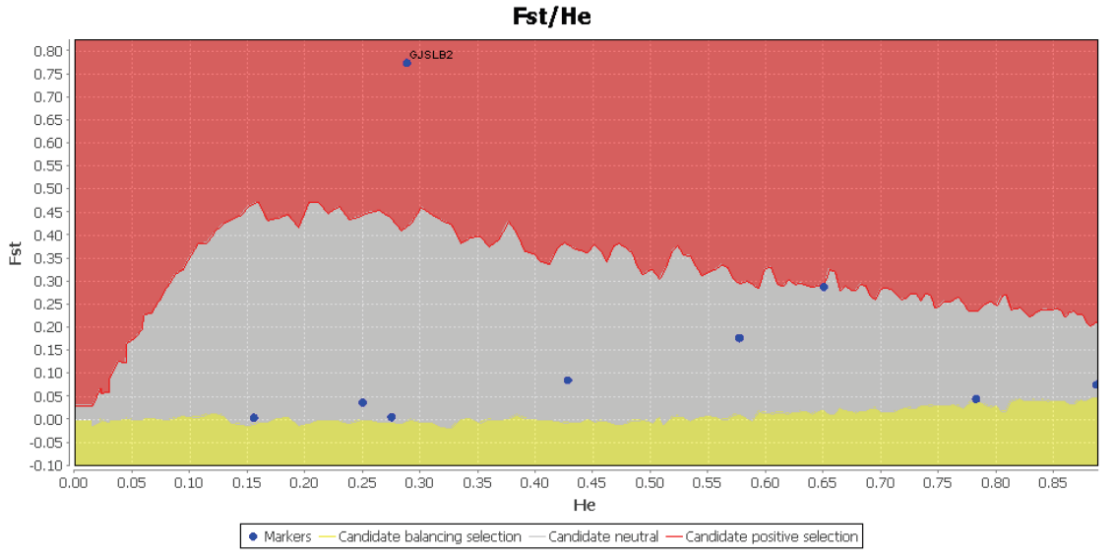



FIGURE 3.7: Fst outlier results under the Infinate Alleles Model (IAM). GJLSB2 is a candidate for positive selection with He=0.289 and Fst=0.774 and 0.288

### 3.4.2 Population structure

#### Mitochondrial DNA data

$\Phi_{ST}$ values indicate a significant level of genetic differentiation between Argentina and all other sites in Australia and New Zealand (Table 3.5). The most prominent difference is between ARG and SWC, with an Fst level of 0.60929. This is followed

closely by 0.60465 between ARG and TAU. The only other statistically significant differences are between SWC and NZL, and SWC and TAU. All significant differences remained after Bonferroni correction. There is a strong isolation-by-distance pattern that can be observed when ARG is included in the analysis ($R^2$=0.117), however this pattern is much weaker once ARG is excluded ($R^2$=0.052) (Figures 3.4 and 3.5).

TABLE 3.5: Pairwise $\Phi_{ST}$ values below the diagonal, with P<0.05 in bold.

| | NZL | ARG | BAR | ESP | PER | SWC | TAU |
|---|---|---|---|---|---|---|---|
| NZL | | 0 | 0.7334 | 0.26953 | 0.28711 | **0.00293** | 0.25684 |
| ARG | 0.57072 | | **0** | **0** | **0** | **0** | **0** |
| BAR | –0.01761 | 0.58801 | | 0.30273 | 0.85742 | 0.08887 | 0.15723 |
| ESP | 0.00865 | 0.56148 | 0.00585 | | 0.41603 | 0.05664 | 0.05273 |
| PER | 0.00485 | 0.58719 | –0.0214 | –0.00186 | | 0.07227 | 0.05859 |
| SWC | 0.07597 | 0.60929 | 0.03138 | 0.03724 | 0.02129 | | **0** |
| TAU | 0.00919 | 0.60465 | 0.02139 | 0.03974 | 0.03539 | 0.15021 | |

All values remained significant after Bonferroni correction. P-values are provided above the diagonal.

The AMOVA data revealed that a significant portion of genetic differentiation, almost 57%, can be attributed to the separation between ARG and Australasia (AUS and NZL). When ARG is excluded from analyses, the percentage of variation that can be attributed to between groups drops significantly and the percentage of variation attributed to within sites increases. The variation between groups within Australasia appears to be greatest between WA and EA at 3.39%. There is no differentiation between EA and NZL which can be seen in the negative %Var values between AUS/NZL and EA/NZL (Table 3.6).

### Microsatellite DNA data

The pairwise FST analyses revealed significant differentiation between ARG and all other sites, however no other pairwise analyses were significantly differentiated, either before or after Bonferroni correction (Table 3.7). These results are concordant with both the G'st and Jost's D results, which revealed significant differentiation between pairwise analyses involving ARG, yet between no others. Jost's D analyses provided the highest values (see Appendix B). The AMOVA analyses were most supportive of a two-population model whereby AUS and NZ represent one population, with ARG representing another (Table 3.8). Under this

model, over 48% of genetic variation can be explained by between group variation. The Mantel test supported a correlation between genetic and geographic distance. However, this result was not statistically significant, nor was it present with ARG removed from the analyses ($R^2$= 0.105798,p=0.077).

The Bayesian Cluster Analysis implemented in STRUCTURE 2.3.4 supported the two-model cluster, with K=2 best describing the allele frequency distribution (Figure 3.8). Under the assumption of K=2, individuals from ARG are assigned to one cluster, with individuals from all other sites assigned to the other cluster. Results were the same in all three runs, however the figure shown, as well as the numbers, are taken from the longest run of 400,000.

The FCA revealed the presence of two main clusters of individuals. As can be seen in Figure 3.9, individuals from ARG all group together excluding one outlier in the top right of the graph which sits on its own, with all other individuals grouping closely in another cluster (Figure 3.10). reveals no pattern in the clustering of all individuals when ARG is excluded.

### 3.4.3 Demographic history

#### Mitochondrial DNA data

High Harpending's Index and SSD values indicate that a sudden expansion model doesn't provide a good fit in the *P.americanus* data, except in the case of NZL with P(SSD)=0.22, and P(R)=0.24. This is concordant with the mismatch distribution data, which deviated significantly from the distribution expected using a model of population expansion (Figures 3.11–3.13). Under a model of population expansion, a unimodal distribution is observed, with the number of pairwise differences between haplotypes directly proportional to the frequency of occurrence. In reality, the distribution of pairwise differences follows an erratic, multi-modal distribution in *P.americanus*. This occurs both when all samples are examined together, and when ARG is analysed separately to Australasia. Additionally, the BSP's lack in support for a recent population expansion (Figures 3.14 and 3.15). In contrast, *P.americanus* appears to have had a more stable historical population size with small increments in population size over the past 115,000 and 200,000 years.

TABLE 3.6: Analysis of molecular variance (AMOVA) of *P.americanus*.

| Regions | Among groups | | | | | Among sites within groups | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | d.f. | Var | % Var | Fct | P-value | d.f. | Var | % Var | Fsc | P-value |
| ARG/WA/EA/NZL | 3 | 0.9956 | 29.2 | 0.29 | 0.07 | 4 | 0.05 | 1.47 | 0.02 | 0.1 |
| Australasia/ARG | 1 | 3.22 | 56.7 | 0.57 | 0.12 | 5 | 0.08 | 1.46 | 0.03 | 0 |
| WA/EA/NZL | 2 | 0.0621 | 2.48 | 0.02 | 0.16 | 4 | 0.05 | 1.98 | 0.02 | 0.1 |
| AUS/NZL | 1 | –0.02 | –0.86 | –0.01 | 0.57 | 5 | 0.1 | 3.88 | 0.04 | 0 |
| EA/NZL | 1 | –0.048 | –1.9 | –0.02 | 1 | 1 | 0.05 | 1.85 | 0.02 | 0.2 |
| WA/NZL | 1 | 0.0447 | 1.79 | 0.02 | 0.4 | 2 | 0.044 | 1.77 | 0.02 | 0.1 |
| WA/EA | 1 | 0.0842 | 3.39 | 0.03 | 0.09 | 3 | 0.05 | 1.9 | 0.02 | 0.1 |

Three apriori geographic groupings: 1) Argentina and Australasia, 2)Argentina, Western Australia, eastern Australia, and New Zealand, 3) Western Australia, eastern Australia, and New Zealand with Argentina excluded from analysis, 4) Australia and New Zealand, 5) eastern Australia and New Zealand, 6) Western Australia and New Zealand, and 7) Western Australia and eastern Australia. D.f. is the degrees of freedom, Var is variance component, % Var is variance component in percentage of total variation.

TABLE 3.7: Pairwise FST ($\theta$) values below the diagonal, significant values after bonferroni correction indicated with *

| | AB | AE | AP | AS | AT | AAR | N |
|---|---|---|---|---|---|---|---|
| AB | | NS | NS | NS | NS | * | NS |
| AE | –0.0004 | | NS | NS | NS | * | NS |
| AP | –0.0068 | 0.0126 | | NS | NS | * | NS |
| AS | 0.0083 | 0.0381 | –0.0041 | | NS | * | NS |
| AT | 0.0114 | 0.0569 | 0.0037 | 0.0122 | | * | NS |
| AAR | 0.4817 | 0.4724 | 0.4592 | 0.4732 | 0.4841 | | * |
| N | –0.0186 | -0.0011 | –0.0124 | 0.0039 | –0.0001 | 0.4684 | |

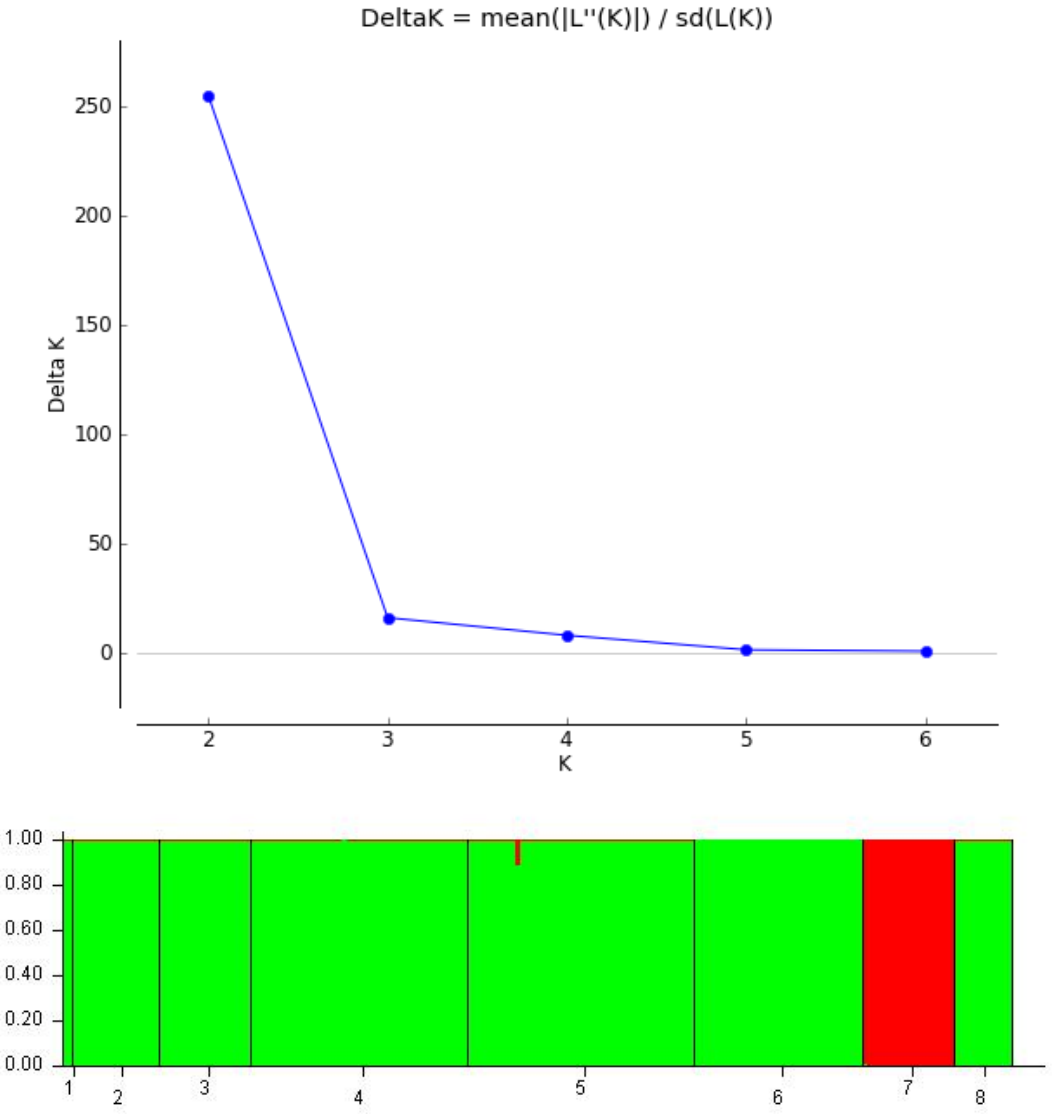



FIGURE 3.8: Results for the Bayesian structure analysis of P.americanus assuming k=2 genetic clusters. Each line represents an individual. Individuals are ordered by site.

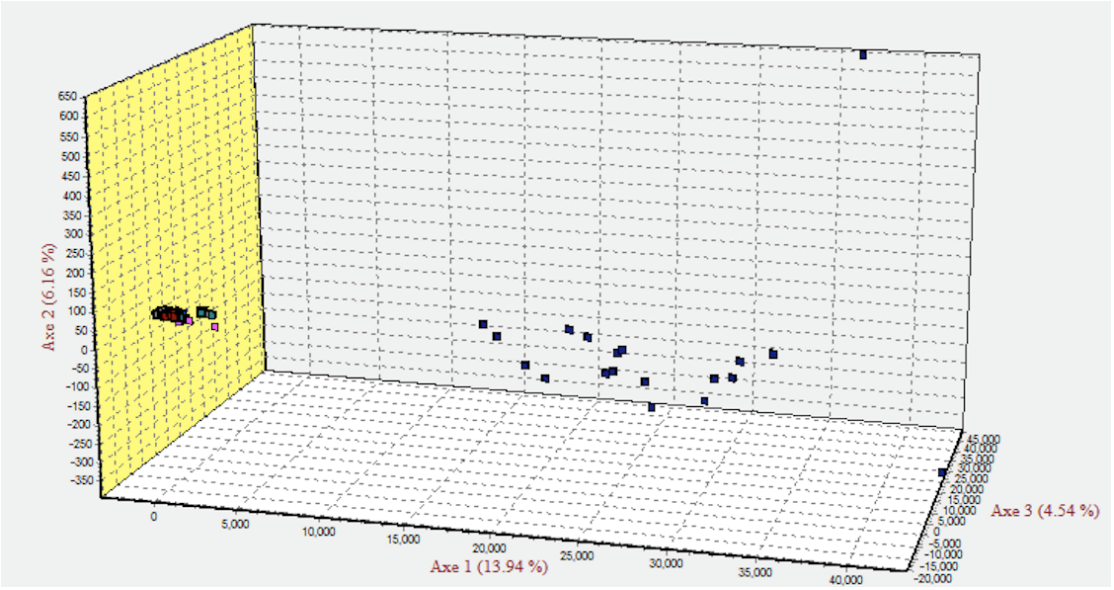



FIGURE 3.9: Factorial components analysis (FCA) of *P.americanus*. The outlier in top right of the graph is an ARG individual.

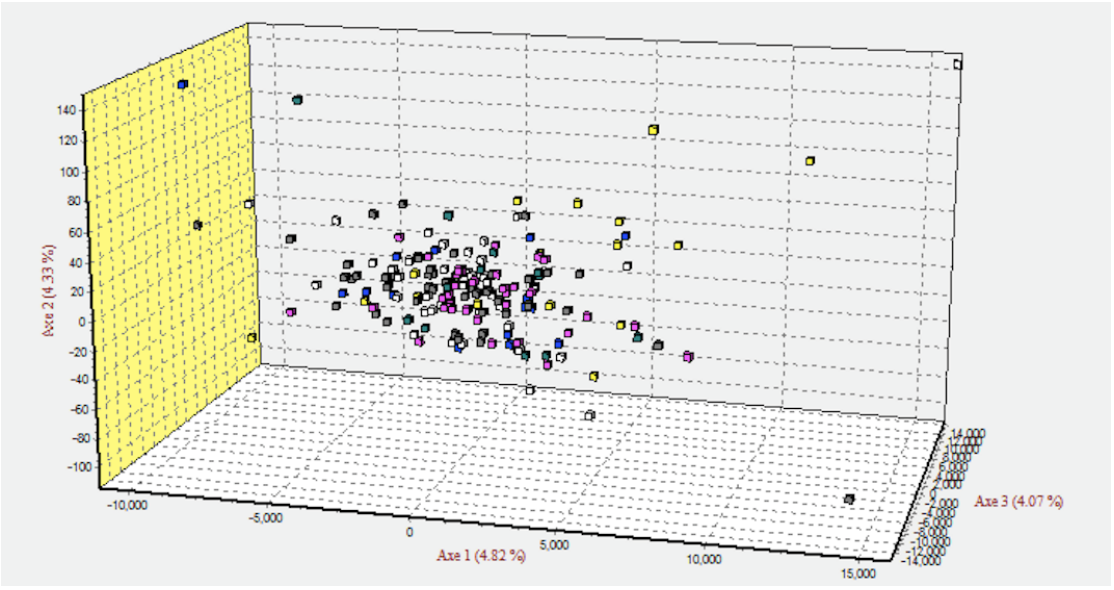



FIGURE 3.10: Factorial components analysis (FCA) of *P.americanus* excluding ARG samples.

The timing of population expansion is estimated to be between 202,070 and 112,261 years ago (ya), using the mutation rates $2.0\text{x}10^{-8}$ and $3.6\text{x}10^{-8}$ respectively. The site that is estimated to have experienced expansion most recently is EA, at 146,818 to 81,565 ya, with WA experiencing expansion the earliest, from 190,328 to 105,737 ya. NZL, which fits the sudden expansion model the best, is estimated to have expanded between 176,868-98,260 ya (Table 3.8).

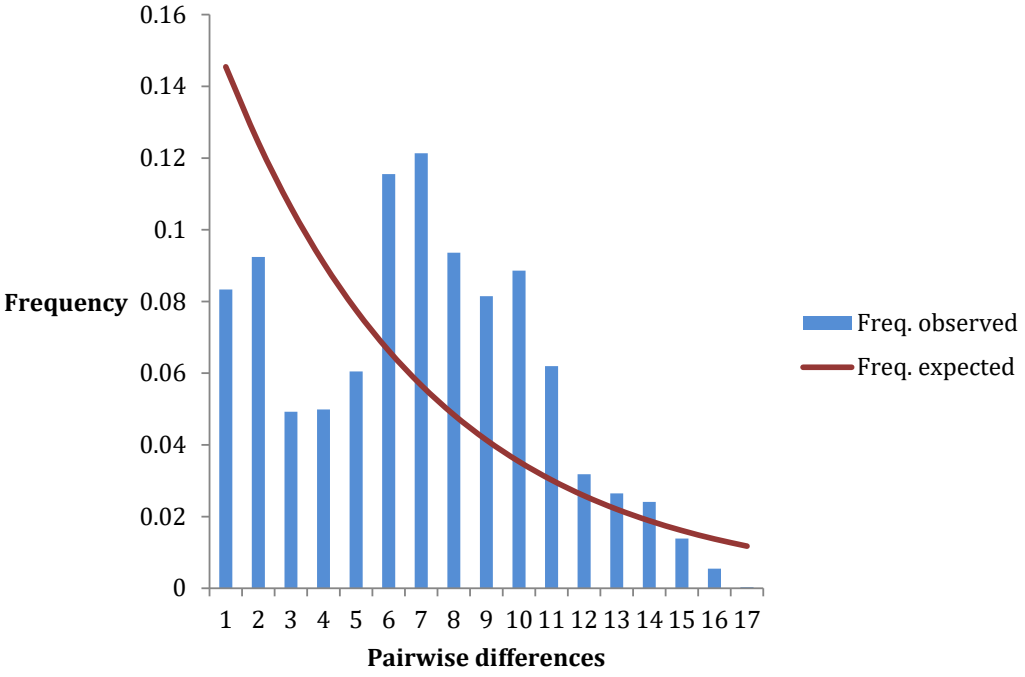



FIGURE 3.11: Mismatch distribution of all pairwise differences for all sites. Freq. expected is under a model of population expansion.

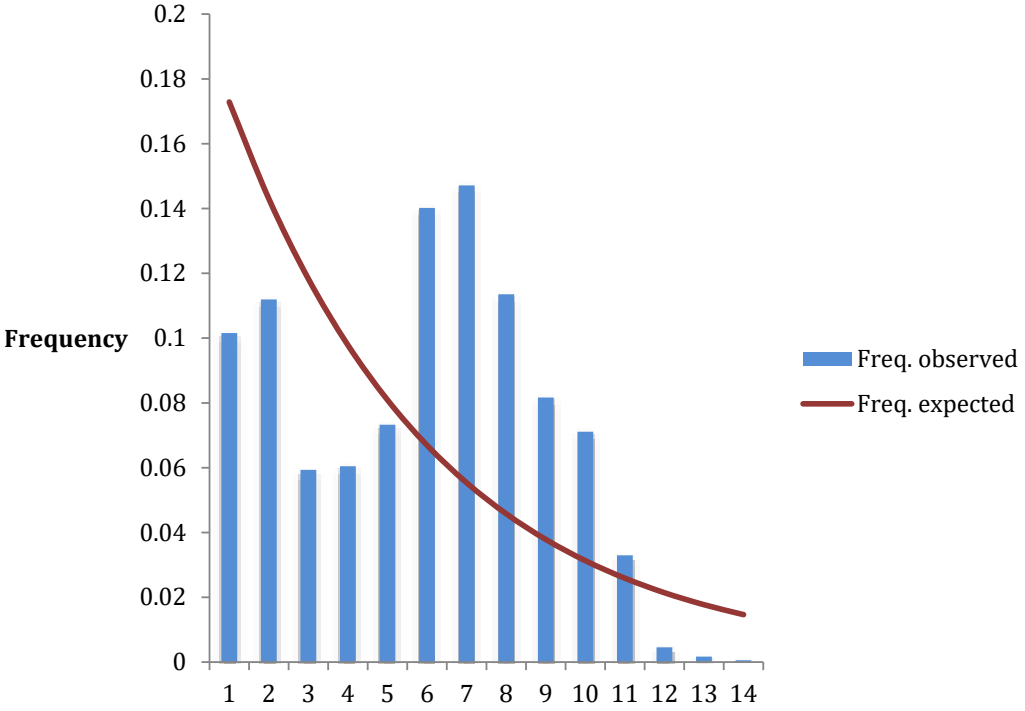



FIGURE 3.12: Mismatch distribution of all pairwise differences for Australasia. Freq. expected is under a model of population expansion.

TABLE 3.8: Microsatellite DNA: Analysis of Molecular Variance (AMOVA) results

| Regions | | ARG/AUSTRALASIA | ARG/AUS/NZ | ARG/WA/EA/NZ | ARG/WA/EA&NZ |
|---|---|---|---|---|---|
| Among groups | DF | 1 | 2 | 3 | 2 |
| | VAR | 0.924 | 0.578 | 0.28 | 0.299 |
| | %VAR | **48.51** | **37.04** | **22.24** | **23.43** |
| | Fct | 0.49 | 0.37 | 0.22 | 0.23 |
| | P | 0.15 | 0.04 | 0.03 | 0.02 |
| Among populations within groups | DF | 5 | 4 | 3 | 4 |
| | VAR | 0.009 | 0.011 | 0.007 | 0.005 |
| | %VAR | **0.49** | **0.68** | **0.53** | **0.4** |
| | Fsc | 0.01 | 0.01 | 0.007 | 0.005 |
| | P | 0.04 | 0.01 | 0.1 | 0.16 |
| Among individuals within populations | DF | 188 | 188 | 188 | 188 |
| | VAR | –0.057 | –0.057 | –0.057 | –0.057 |
| | %VAR | **–2.98** | **–3.64** | **–4.51** | **–4.45** |
| | Fis | –0.06 | –0.06 | –0.06 | –0.06 |
| | P | 0.99 | 0.99 | 0.1 | 0.99 |
| Within individuals | DF | 195 | 195 | 195 | 195 |
| | VAR | 1.028 | 1.028 | 1.028 | 1.03 |
| | %VAR | **53.99** | **65.92** | **81.75** | **80.62** |
| | Fit | 0.46 | 0.34 | 0.18 | 0.19 |
| | P | 0 | 0 | 0 | 0 |

for four groupings: 1) Two regions, ARG and Australasia (AUS and NZL) grouped separately; 2) Three regions, ARG, AUS (WA and EA) and NZL grouped separately; 3) Four regions, ARG, WA, EA, and NZL grouped separately; and 4) Three regions, with ARG, WA and the rest (EA and NZ) grouped separately.

TABLE 3.9: Estimation of population expansion using Fu's Fs, and mismatch distribution analyses: sum of squared deviation (SSD), Harpending's raggedness index,Tau ($\tau$) , population size before expansion ($\theta$o), population size after expansion ($\theta$1), and time since population expansion occurred (t) with P-values and confidence intervals in brackets.

| | Fu's Fs (P) | Tajima's D (P) | SSD (P) | Raggedness (P) | t (95% CI) | Oo (95% CI) | O1 (95% CI) | t |
|---|---|---|---|---|---|---|---|---|
| All sites | –24.85966 (<0.00) | –1.14065 (0.11) | 0.0024 (0.88) | 0.0063 (0.94) | 8.002 (2.5–12.0) | 0.025 (0.00–2.95) | 13.15 (8.55-122.68) | 202,070-112,261 |
| ARG | -5.90834 (<0.00) | –0.35083 (0.40) | 0.0100 (0.60) | 0.0212 (0.84) | 6.080 (2.12–9.54) | 0.00 (0.00–3.19) | 12.91 (6.90-99999.0) | 153,535-85,297 |
| WA | –17.82081 (<0.00) | –0.80493 (0.23) | 0.0101 (0.65) | 0.0182 (0.71) | 7.537 (1.50–12.68) | 0.00 (0.00–2.41) | 8.73 (4.79-99999.0) | 190,328-105,737 |
| EA | –6.55037 (0.01) | –1.07647 (0.14) | 0.0058 (0.55) | 0.0169 (0.64) | 5.814 (2.26–8.77) | 0.00 (0.00–2.27) | 18.46 (9.99-99999.0) | 146,818-81,565 |
| NZL | –3.02851 (0.10) | –0.79473 (0.23) | 0.0136 (0.22) | 0.0383 (0.24) | 7.004 (2.93-10.12) | 0.00 (0.00–3.40) | 25.31 (12.76-99999.0) | 176,868-98,260 |

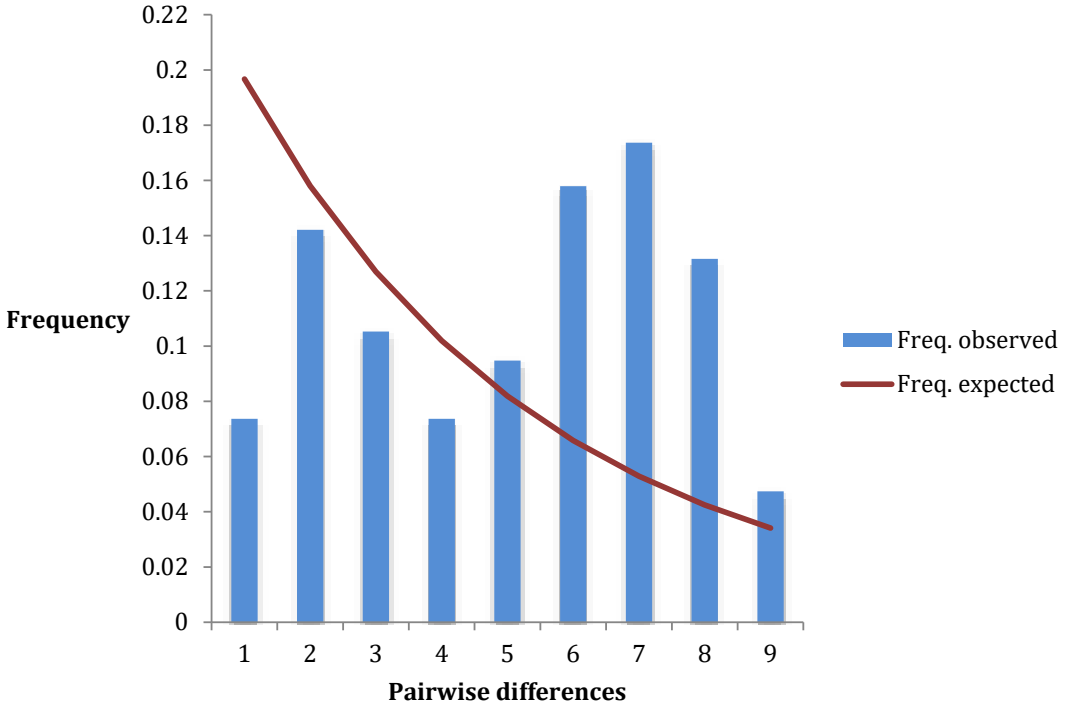



FIGURE 3.13: Mismatch distribution of all pairwise differences for Argentina. Freq. expected is under a model of population expansion.

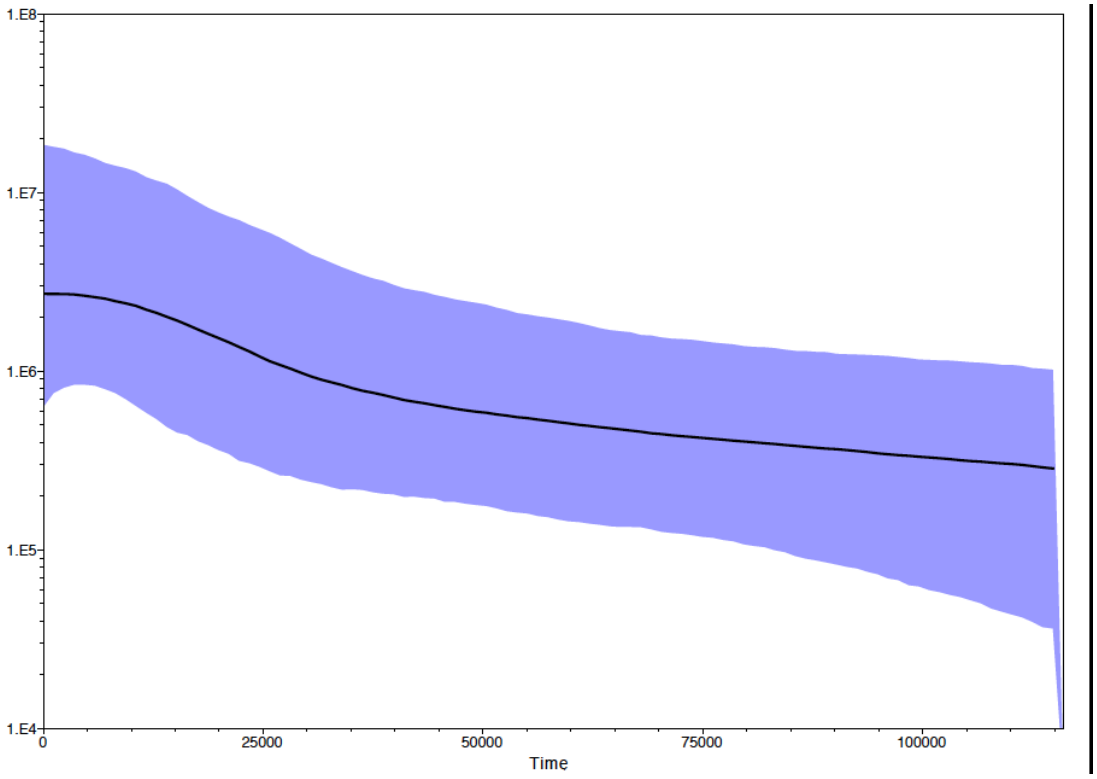



FIGURE 3.14: Bayesian skyline plot with $\mu$=2.0x10$^{-8}$. Time=time before present. Y axis=effective population size.

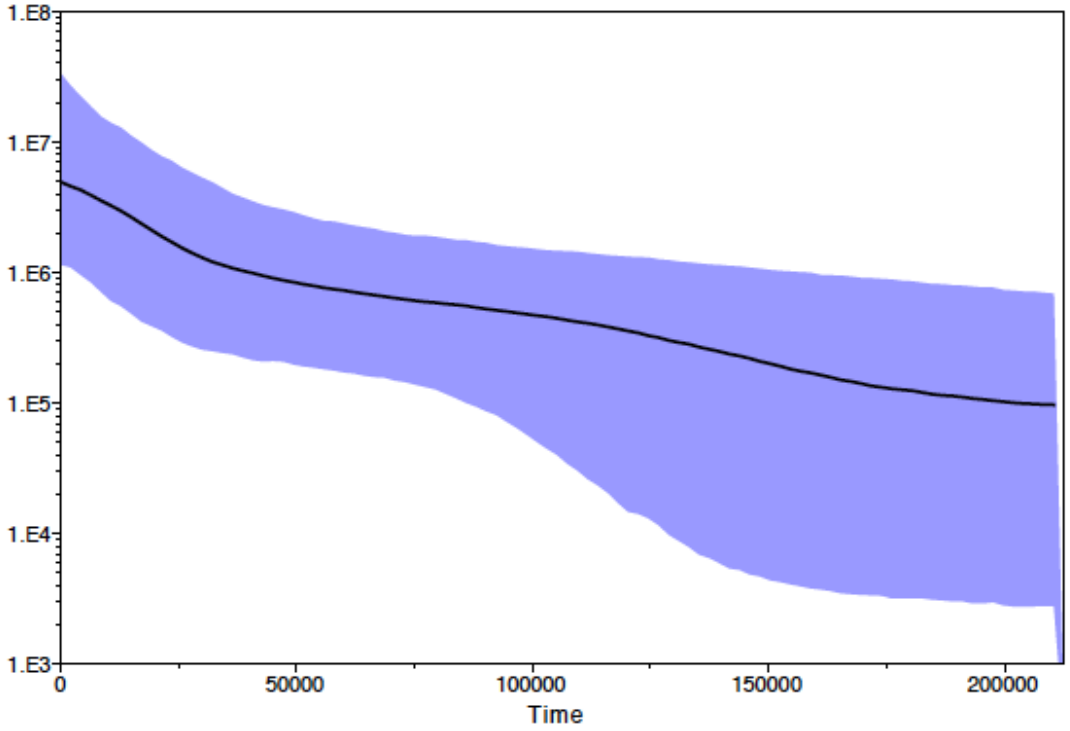



FIGURE 3.15: Bayesian skyline plot with $\mu$=3.6x10$^{-8}$. Time=time before present. Y axis=effective population size.

## 3.5 Discussion

### 3.5.1 Genetic diversity

The observed levels of haplotype diversity were high across all sites, indicative of large historical population sizes (Varela et al. 2013). This result is fairly typical for a marine species, which tend to have larger populations than similar freshwater species. This is partially due to their wider geographic ranges in environments that do not tend to limit gene flow, and this creates large effective population sizes which reduces their susceptibility to strong drift (DeWoody & Avise 2000). Larger populations are able to hold more haplotypes, with a larger population size increasing the potential for new mutations and reducing the strength of drift, leading to higher levels of diversity (Shaw et al. 1999).

The mtDNA and microsatellite DNA results support a two-population model for *P. americanus* in the Southern Hemisphere. The level of differentiation between Argentina and all other locations suggests that Argentina is an isolated and demographically distinct group. Ball et al. (2000) also reported on two southern *P.americanus* populations based on microsatellite DNA analysis: A Brazilian, or South Western Atlantic stock; and a Pacific stock comprising of individuals from eastern Australia and New Zealand. There was too much missing data at the microsatellite loci PAM010, PAM025, and GH0OIK in the ARG samples to include these loci in further analyses of the Argentina site. However, the only null allele detected in Australasia was found at one locus in the Perth samples. Null alleles are important in population genetic studies as, although they may represent a technical problem, they often represent a mutation at a primer site (Chapuis & Estoup 2007).

The present study found some evidence of differentiation between Western Australia and the rest of Australasia; however that result only appears in the pairwise comparisons with South West Capes when using the mtDNA data. This potentially reflects a need for greater sampling, as was determined from the rarefaction curves. In contrast to the results from this study, Sedberry et al. (1996) was unable to detect any significant differentiation within the Southern Hemisphere. Differentiation within Australasia in deep-sea teleosts has previously been recorded. For example, Elliot et al. (1998), using allozyme markers, detected evidence for three geographic stocks of spiky oreo (*Neocyttus rhomboidalis*) in Australasian waters:

1) East coast of New Zealand, 2) Western and South Australia, and 3) western Tasmania to New South Wales. Future *P. americanus* genetic studies that included samples from the western Indian Ocean could add an important site, especially because it has been speculated that a third species of *Polyprion* exists off South Africa, with distinct morphological differences and 'alleles' observed (Ball et al. 2000, Robert 1986). In contrast to the findings here and with Elliot et al. (1998), Varela et al. (2013) was unable to detect significant levels of differentiation amongst Southern Hemisphere populations of orange roughy (*Haplostethus atlanticus*) using mitochondrial DNA. This is surprising given the geographical distances between the sites sampled (Australasia, Namibia, and Chile), however panmictic structuring possibly arises from their ability for dispersion as adults and their longevity (Varela et al. 2013).

As mentioned in chapters 1 and 2, population structure can generally be placed into three continuous categories: 1) Panmixia (no differentiation), in which genetic homogeneity prevails over the geographic region considered; 2) Isolation-by-distance (continuous change), in which the genetic composition changes continuously over space; and 3) Distinct populations, where gene flow is small enough to permit genetic divergence among closely related populations (Laikre et al. 2005). *P.americanus* structuring appears to lie along the continuum closer to panmixia at smaller scales, with some differentiation at larger scales. The Mantel test did not support a model of isolation-by-distance.

### Argentina

*P.americanus* experience pan-oceanic mixing in their juvenile years through their association with flotsams. It isn't until they mature at four years that they migrate downwards and potentially display greater homing behaviour. In the Northern Hemisphere this has resulted in the production of one large panmictic population with juveniles moving westward with the major currents before migrating deeper off the Azores (Sedberry et al. 1996). While it appears that *P.americanus* are genetically one population in eastern Australia and New Zealand, *P.americanus*, for whatever reason, are not able to utilise the dominant circumpolar flow in the Southern Hemisphere which largely connects the Atlantic, Pacific, and Indian Oceans (Gordon 1971, Ball et al. 2000). This is potentially due to the summer surface temperatures ($<10°C$) in the Antarctic Circumpolar Current or West Wind

Drift, which are cooler than waters that *Polyprion* have previously been observed in (Sedberry et al. 1999, Ball et al. 2000). Argentina represents a distinct genetic stock, supported by both the mtDNA and microsatellite DNA data. All $\Phi_{ST}$, Fst, Jost's D, and G'st pairwise comparisons involving ARG were significantly differentiated. Further, AMOVA, STRUCTURE analysis, and FCA analysis supported the separation of Argentina from Australasia.

**Australasia**

Using indirect genetic approaches, populations are defined by a number of migrants, not a proportion. In larger populations this means that demographically uncoupled populations may be left unidentified purely because the number of migrants may be too high to detect differentiation. For example, 120 migrants would have a larger ecological effect on a population of 1000 than it would on a population of 100,000 (Hauser & Carvalho 2008, Ward 2000). In this way, microsatellite DNA markers can provide limited value for detecting differentiation. A population size one quarter that of microsatellite DNA markers means that mtDNA may reach equilibrium sooner and provide detectable differentiation where nuclear markers cannot (Hauser & Carvalho 2008, Vigliola et al. 2007). On the other hand, the high mutation rate of microsatellite DNA can compensate for this. With this in mind, there are two possible origins for the discordance between the mtDNA and microsatellite DNA markers in the Australasian samples:

1. An excess of heterozygotes in the microsatellite DNA markers has reduced the Fst values in the current study. This is common in species with large effective population sizes that can accumulate more mutations than smaller populations. Larger populations also take longer to reach mutation-drift equilibrium. Populations that have recently undergone bottleneck contraction can also have inflated levels of heterozygosity. However, the observed heterozygosity for South West Capes does not differ significantly from expected levels (Ho= 16.6, He=17.1) (Holsinger & Weir 2009). Conversely, microsatellite DNA tends to accumulate mutations at a faster rate than mtDNA, which can lead to high levels of homoplasy in a data set.

2. A second possibility, and one that seems more plausable, is that the differentiation observed in the mtDNA reflects historical patterns, but the contemporary populations are in secondary contact . This is one reason why

multiple loci and marker types are preferable in population genetic studies. For example, mtDNA and microsatellite DNA were used to detect variation at different geographical scales in Atlantic herring stocks of *C.harengus*. While the microsatellite DNA detected variation at a finer scale within the Atlantic, the mtDNA was able to detect variation between the Pacific and Atlantic Ocean basins that was undetected using microsatellite DNA markers (Shaw et al. 1999).

If genetic differentiation detected in the data set does faithfully match the contemporary structure of the populations, it might suggest that the unique current systems in Australia are more important for maintaining the genetic connectivity of *P.americanus* than geographic proximity as proposed for *P.oxygeneios* in Chaper 2 (Aboim et al. 2005). While the East Australian Current is effective at transporting pelagic larvae offshore for most parts of the year (Condie et al. 2011) the Leeuwin current entrains larvae, carrying them south. While *P.americanus* juveniles are not passive larvae, the flotsams that they are associated with are completely passive and are transported via currents in the same way as passive larvae. Thus, the flow of currents around Australia could both promote mixing between eastern Australia and New Zealand, and limit mixing between Western and eastern Australia, with a potential source-sink population scenario occurring (Hauser & Carvalho 2008). There are examples of current systems controlling gene flow among pelagic fish populations. For example, Blue Mackerel in Australia potentially utilise the East Australian Current system to disperse into more productive waters off of New Zealand, consistent with the genetic homogeneity observed between these areas (Condie et al. 2011, Scoles et al. 1998). Blue Mackerel are different in that their juvenile phase is relatively short when compared to *Polyprion* (29 days or less), however their juveniles are transported passively by current systems (Scoles et al. 1998). There are deep-sea examples also. The Benguela Current system in the southeastern Atlantic likely forms a historical and contemporary barrier to gene flow in the deep-sea bluemouth *(Atractoscion aequidens)*, as seen using both mtDNA and microsatellite DNA markers (Aboim et al. 2005).

**Demographics**

NZL is the only site that appears to fit a model of sudden population expansion. All other sites fit a model of constant population size, common in deep-water species which tend to be less affected by environmental changes connected to glacial cycles (Gaithier et al. 2011). In contrast, *P.oxygeneios* in New Zealand is reported to have high levels of both haplotype and nucleotide diversity, and does not fit a model of population expansion (Lane et al. 2016). After losing genetic variation through a bottleneck event, a population will accumulate haplotype diversity faster than nucleotide diversity (Lane et al. 2016, Varela et al. 2013). When a population reaches mutation-drift balance, the number of haplotypes plateaus, and the ratio of haplotype diversity to nucleotide diversity decreases (Lane et al. 2016). Thus, by plotting haplotype diversity against nucleotide diversity, the post bottleneck stage of a population can be seen (Lane et al. 2016). Although population size expansion in NZL has possibly followed a bottleneck effect, further sampling from around NZL is needed to create a more holistic picture of the area. There was no evidence for heterozygotic excess in NZL, common in populations that have recently undergone population contraction. There was also no statistical support for the negative Fu's F value either ($p>0.05$). However, heterozygosity measures are less sensitive to the effects of Bottleneck than other measures, such as allelic diversity (Allendorf et al. 2007). The figures for population size expansion in NZL fall within the Pleistocene period (176,868-98,260 ya) which had a large effect on coastal species in New Zealand. There is evidence that *P.oxygeneios* once inhabited shallower waters in New Zealand (Maxwell 2011). This could possibly apply to *P.americanus* also. Of potential importance to note is that false differentiation has been detected in populations that have recently undergone bottleneck effects, as these populations are not always in drift-equilibrium (Hauser & Carvalho 2008).

The high number of unsampled mutational steps in the haplotype network supports the existence of a stable *Polyprion* population in the Southern Hemisphere. A population that has recently undergone expansion will typically display a star-shape pattern (Xiao et al., 2010). A constant, and stable population is further supported by the BSP's, which do not support the existence of a single expansion event in the last 100,000- 200,000 years using the mutation rates of $3.6\text{x}10^{-8}$ and $2.0\ \text{x}10^{-8}$. The data from the mismatch distributions estimate a population expansion time of ∼200,000–115,000 ya, which potentially supports the BSP's if

the earlier time of 200,000 ya is appropriate. The timing of a population expansion is however hugely dependent on the chosen mutation rate (Varela et al. 2013), which has not been calibrated for *Polyprion*.

### Implications for fisheries management

Genetic differentiation seen in mtDNA is thought to reflect historical patterns to gene flow, however management decisions are usually based on current levels of exploitation and migration (Hauser & Carvalho 2008). From the results seen here in the msat DNA, it could be assumed that the physical boundaries separating South West Capes from other areas have disappeared, with secondary contact having since occurred. However, low levels of genetic differentiation observed in msatDNA should be considered carefully. Microsatellite DNA is more prone to effects of homoplasy and requires large sample sizes. Further, a lack of genetic differentiation does not necessarily equate to demographic connectivity (Lowe et al. 2010). With more biological and genetic data, the origin for the observable differentiation could be studied, which would be hugely beneficial for policy makers. Temporal and spatial replicates could confirm the presence of weak genetic differentiation in the msat DNA data (Hauser & Carvalho 2008).

### Summary

In summary, the Southern Hemisphere appears to consist of two discrete populations of *P.americanus*. There is detectable variation between Western Australia and the rest of Australasia, however the demographic implications, namely whether this reflects historical or contemporary gene flow, of this differentiation are unknown. Further, greater sampling from sites in New Zealand is needed to determine whether Australasia truly represents one biological population. If *P.americanus* has a similar structuring to *P.oxygeneios* then there could possibly be more than one population present in New Zealand. Future sampling from the Western side of the Indian Ocean would also be beneficial, especially given the speculation of a third *Polyprion* species (Ball et al. 2000).

# Chapter 4

# Discussion

## 4.1 General

This study was the first to use both mtDNA and microsatellite DNA markers to examine the genetic connectivity of both *Polyprion* species in the Southern Hemisphere. The results of the analyses of genetic data from both species means the null hypothesis of overall panmixia was rejected, with significant levels of genetic differentiation observed at both the small and the large scales. The findings of this thesis research advanced on the findings made in previous studies of *P. americanus*. Ball et al. (2000) detected significant levels of differentiation for *P. americanus* based on an analysis of microsatellite DNA markers. A prior study of *P. oxygeneios* (Lane et al. 2016) lacked a large set of samples from a broad ranges of locations. The finding of genetic differentiation between Argentina, Cape Town, and Hokitika with sites in Australasia most likely supports a stock model of demographically uncoupled populations, which is important information for fishery management.

### 4.1.1 Genetic diversity and population structure

Both *Polyprion* species have highly mobile adults and juveniles with a long pelagic phase. These characteristics led to the assumption that *Polyprion* would be more or less one large homogenous population in the Southern Hemisphere, similar to their Northern Hemisphere counterparts (Ball et al. 2000, Sedberry et al. 1999). In contrast to expectations, differentiation exists between *P. oxygeneios* at the sample site near South Africa and sites in Australasia. Hokitika in the New Zealand EEZ was differentiated from many other sampled locations based on the msatDNA data, and the Chatham Islands sample location was significantly differentiated based on the mtDNA data only. Mitochondrial DNA analyses revealed differentiation between *P. americanus* in Argentina and all other sampled sites in the Southern Hemisphere. Although Ball et al. (2000) reported a similar pattern of genetic structure, their findings were only based on the use of microsatellite DNA markers. In addition, that study lacked samples from locations in Western Australia, which was an area that this thesis research did sample and the mtDNA results showed to be differentiated from other Australasian locations. With the use of genetic markers, differentiation is commonly detected between sites that were previously assumed to be homogenous (Hedgecock et al. 2007). For example, Atlantic Cod *(Gadus morhua)* exhibit differentiation at relatively small spatial

scales (<3000km), despite being capable of long distance migration with a broad geographical distribution. It is thought that oceanographic features (i.e. gyre-like eddies) 'trapping' cod are more important in determining patterns of genetic connectivity than distance (Ruzzante et al. 1998). In contrast, Varela et al. (2012) recently reported Orange Roughy in Australasian waters to be largely homogenous despite there being large distances between sample sites. These results highlight the importance of studying each species on a case by case basis (Hauser & Seeb 2008).

Previously published research that used mtDNA was unable to detect a significant level of differentiation among *P. americanus* in the Southern Hemisphere (Ball et al. 2000, Sedberry et al. 1996). The large population sizes that are typical for marine fish mean that to detect low levels of differentiation, a large number of samples are required to obtain adequate statistical power, compared to other species, such as freshwater fish that tend to be more highly structured due to limits in their dispersal opportunities. It is possible that the sample sizes reported in Ball et al. (2000) of 28 samples from Brazil and 30 samples from New Zealand were too low to detect differentiation. However, the sample sizes from many locations in the present study were comparable to Ball et al. (2000). Sedberry et al. (1999) proposed the possibility of a distinct *P. oxygeneios* stock in northern New Zealand, however based on the genetic data presently available, there is no evidence to support that type of structure. A more substantial sample from New Zealand sites might provide a more definitive statistical test of Sedberry's stock hypothesis.

New alleles are brought about by mutations which are either lost, or become fixed in a population through selection and random genetic drift. The movement of individuals acts to homogenise populations, and it is the balance between mutations and migration that results in the genetic structure that can be observed in a population. In general terms, population genetic structure can be placed into three categories: 1) Panmixia (no differentiation), in which genetic homogeneity prevails over the geographic region considered; 2) Isolation-by-distance (continuous change), in which the genetic composition changes continuously over space; and 3) Distinct populations, where gene flow is small enough to permit genetic divergence among closely related populations (Laikre et al. 2005). Distinct populations and isolation-by-distance is common in species where proximity is an important factor for mating success or opportunity. The relative genetic homogeneity of *Polyprion* suggests that migration over large geographic distances is relatively common, and

supports assumptions that both species travel with currents as juveniles in the Southern Hemisphere (Ball et al. 2000, Lowe et al. 2010).

However, there are several characteristics of *Polyprion* and its environment that might limit gene flow. First is that reproduction might occur in aggregations of fish that form at scales much smaller than the potential dispersal range of *Polyprion* (Broquet et al. 2012) and that they tend to migrate in more-or-less stable groups between spawning and feeding areas. A strongly cohesive reproductive unit might have a small genetic effective population size, which would subject it to strong genetic drift. However, very little is know about their reproductive behaviour, breeding locations-, and the levels of relatedness within a reproductive aggregation (Lane et al. 2016, Wakefield et al. 2010; 2013). Second, post-settlement selection might be an important factor in a heterogenous environment driving selection for locally adapted variants (Broquet et al. 2012). Temporal genetic studies with high resolution genomic markers would be needed to determine the stability of the genetic variation through time and identify candidate genes for selection. There is also the possibility that the genetic differences observed in contemporary studies are a reflection of historic barriers to gene flow, but modern populations have now come back into secondary contact (Lane et al. 2016). With differentiation found only from the analysis of the mtDNA data, and not using the msatDNA data, it lends credibility to the idea there was in the past a lack of gene flow between *P. americanus* in South West Capes and other locatons.

Of all the potential physical barriers, oceanic currents are cited as a major factor connecting the northern *P. americanus* (Sedberry et al. 1999). It has already been discussed as a possible mechanism for the differentiation observed between South America and other southern sites (Ball et al. 2000, Lane et al. 2016). Although there are no obvious physical barriers to *P. oxygeneios* between Hokitika and other sites (Ross et al. 2009, Lane et al. 2016), there is the strong potential for oceanic currents to act as barriers between Cape Town and other sites, and Western Australia with the rest of Australasia (Condie et al. 2011, Beckley et al. 2009). The Benguela and Agulhas currents are the predominant currents off South Africa, however neither is directly connected to Western Australia or other parts of Australasia. Further, while the East Australian Current is effective at transporting pelagic larvae offshore for most parts of the year (Condie et al. 2011) the Leeuwin current entrains larvae, and transports them south (Molony et al. 2011). *P. americanus* juveniles are not passive larvae, however they have

a strong association with flotsams which are completely passive and transported via surface currents in similar ways that currents might transport passive larvae. Thus, the flow of currents around Australia could both promote mixing between eastern Australia and New Zealand, and limit mixing between western Australia and the rest of Australasia (Condie et al. 2011). There are examples available for other commercial demersal fish species showing the current systems in Australia to affect recruitment (Condie et al. 2011), including King George Whiting (Fowler et al. 2000, Jenkins et al. 2000) and Blue Grenadier (Bruce et al. 2001).

### 4.1.2 *Polyprion* taxonomy

The results from this thesis research were not inconsistent with the idea that there are more than two species of *Polyprion* in the Southern Hemisphere. Firstly, the level of differentiation observed between *P. oxygeneios* in South Africa and Australasia (Fst>0.16 at all sites) was consistent with the findings reported by both Robert (1986), who recorded distinctive morphological features in what is thought to be four *P.americanus* individuals, and Ball et al. (2000) who reported that a third species of *Polyprion* exists off South Africa. Secondly, the results support the findings that *P. moene* and *P. oxygeneios* are the only valid species in New Zealand's EEZ (Ball et al. 2000). *P. americanus* in New Zealand was previously recorded as a separate species, *P. moene*, but was synonymised with *P. americanus* by Robert (1986). However, the levels of differentiation within *P. americanus* were sometimes greater than between the two species of *Polyprion* in Ball et al. (2000). For example, the pairwise difference between *P. americanus* in New Zealand and Brazil using the diversity measure $(\delta\mu)^2$ is 81.93 in comparison to 31.90 between Brazil and *P. oxygeneios*. The *P. americanus* samples from the Australian and New Zealand sample sites in the present study were largely undifferentiated which supports the idea that both sites contained the same species. In which case, *P. moene* would be present throughout Australasia.

### 4.1.3 Demographics

The Pleistocene period (>10,000 years before present) was punctuated by a series of large glacial-interglacial changes, known to coincide with large bottleneck and expansion events in many coastal species. Deep-water species, such as *Polyprion*,

may be less affected by fluctuating environmental conditions associated with glacial cycles, however there is evidence that *P. oxygeneios* once inhabited shallower waters in New Zealand (Maxwell et al. 2011). During this time period, *P. oxygeneios* appears to have experienced population expansion, expanding earlier in Australia than New Zealand. *P. americanus* in New Zealand also fits a model of population expansion, however the estimated time period is much earlier at 176,868-98,260 years ago (ya) in contrast to *P. oxygeneios*, which appears to have expanded 18,000-33,000 ya. *P. americanus* in the rest of the Southern Hemisphere does not fit a model of population expansion. During the last glacial maximum, sea levels in New Zealand were much lower, exposing more land than is currently seen today, and joining the three main islands. It has previously been speculated that Cook Strait in New Zealand could be a possible spawning site for *P. oxygeneios* (Paul 2002). This site would not have been underwater during the last glacial maximum, meaning another spawning site in New Zealand or elsewhere probably existed. With Cook strait cut off, *P. oxygeneios* in Hokitika along the west coast of New Zealand potentially spawned in a separate location to *P. oxygeneios* elsewhere in New Zealand, fitting with the hypothesis that the observable differentiation between this site and others reflects historical differentiation. Although it is speculative, it is probable that *P. americanus* dispersed into the Southern Hemisphere before *P. oxygeneios* speciated, supported by an earlier expansion time, and a more stable population history.

### 4.1.4 Management

While it is typically harmless to over-split a stock, harvesting a group of individuals that belong to separate populations as if they are a single group (termed a mixed-stock fishery), may result in the loss of genetic diversity in one or more of the sub-populations (Allendorf et al. 2008). Currently the New Zealand HPB (Hapuku-Bass) fishery is separated into eight stocks, however the HPB stock boundaries might be over-split in New Zealand, with the results here suggesting that *P. oxygeneios* in New Zealand likely represent one panmictic population, with some fine scale differentiation observable. While additional markers may reveal further levels of differentiation within the New Zealand EEZ, the data presented here provides genetic support for the current management of Hokitika as a separate stock to the rest of New Zealand only. The management areas in New

Zealand are divided up based on geographic features and concentrations of fishing effort. The stock boundaries themselves differ from the boundaries used to determine regional yields, or the tonnage of fish extracted from an area (Paul 2002). Cook Strait, for example, is situated on the borders of three stocks, HPB 2, HPB 7, and HPB8. While a maximum sustainable yield (MSY) is used to set the respective total allowable commercial catches (TACC) of each stock in order to reduce over-exploitation, it is currently unknown how relevant stock boundaries are to the natural population boundaries of *P. americanus* (MPI Science Group 2015). The inclusion of *P. americanus* samples from throughout New Zealand in future analyses is necessary to determine whether the stock boundaries in the HPB fishery reflect natural population boundaries of both *Polyprion* species.

In south-western Australia, annual commercial catches of *Polyprion* have increased since the first recorded state commercial landings in 1986 (Wakefield et al. 2010). Compared to the 2179 t caught in the New Zealand fishery (2002), the south-western Australian catch is relatively small at 18 t in 2010. The smaller Australian fisheries are partially due to nutrient poor waters that do not support high productivity (Paul 2002; Molony et al. 2011). Despite the low recordings and the risk status reported as moderate (Fletcher & Santoro 2015), the current resource has previously been suggested to be fully exploited in south-western Australia (Wakefield et al. 2010). If the differentiation observed between *P. americanus* in South West Capes and other sites reflects contemporary differentiation, this suggests either a physical barrier between this site and others, or potentially the effects of fishing in the area. Temporal and spatial replicates could confirm the presence of weak genetic differentiation (Hauser et al. 2008). Differentiation at mtDNA loci is however usually considered to reflect historical barriers to gene flow which is slower to 'evolve' than nuclear loci (Schloetterer et al. 2004). Further, the differentiation was not observed in *P.oxygeneios* either, which might be expected given the ecological similarities between both species.

## 4.2 Summary

Increasing fishing exploration into deep-water coupled with a higher inherent vulnerability of *Polyprion* species creates a strong need to better understand the

population connectivity. *Polyprion* are relatively long lived and have a low reproductive output (K-selected biology), rendering them more vulnerable to overexploitation than faster growing fish species with a greater reproductive output (R-selected species). In general, *Polyprion* can be considered panmictic, with differentiation observable between ocean basins. Whether this differentiation reflects speciation or strong population differentiation is uncertain, however it is probable that currents in the Southern Hemisphere play an important role in connectivity patterns for both *Polyprion* species. In general, *P. americanus* appears to be a more stable population than *P. oxygeneios*, potentially reflecting a greater amount of time present in the Southern Hemisphere. The expansion times for both species fall within the Pleistocene period, marked by many changes to coastal topography and thus genetic connectivity patterns in many marine species (Maxwell et al. 2011). While in New Zealand the management areas of HPB incorporate the fine scale differentiation observed within *P. oxygeneios*, it is uncertain whether this is applicable for *P. americanus* also. The south-western Australian stocks of *Polyprion* may already be overfished, and so the potential differentiation observed between South West Capes and other sites should be viewed with caution.

# Appendix A

# Additional information for *P. oxygeneios*

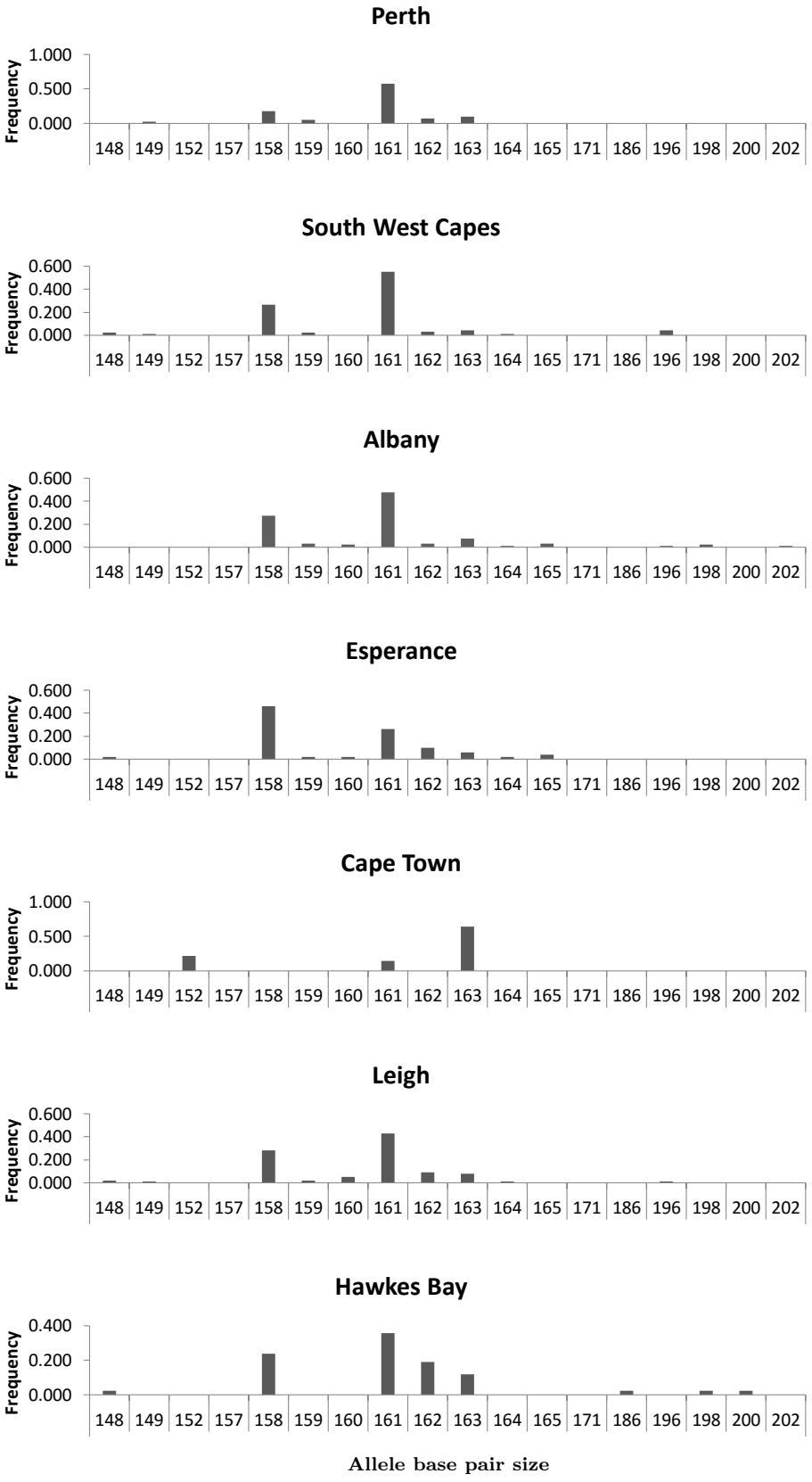



FIGURE A.1: Allelic Frequency Distribution for the microsatellite locus GG0OQ6A. Allele sizes are given in base pairs of PCR products.

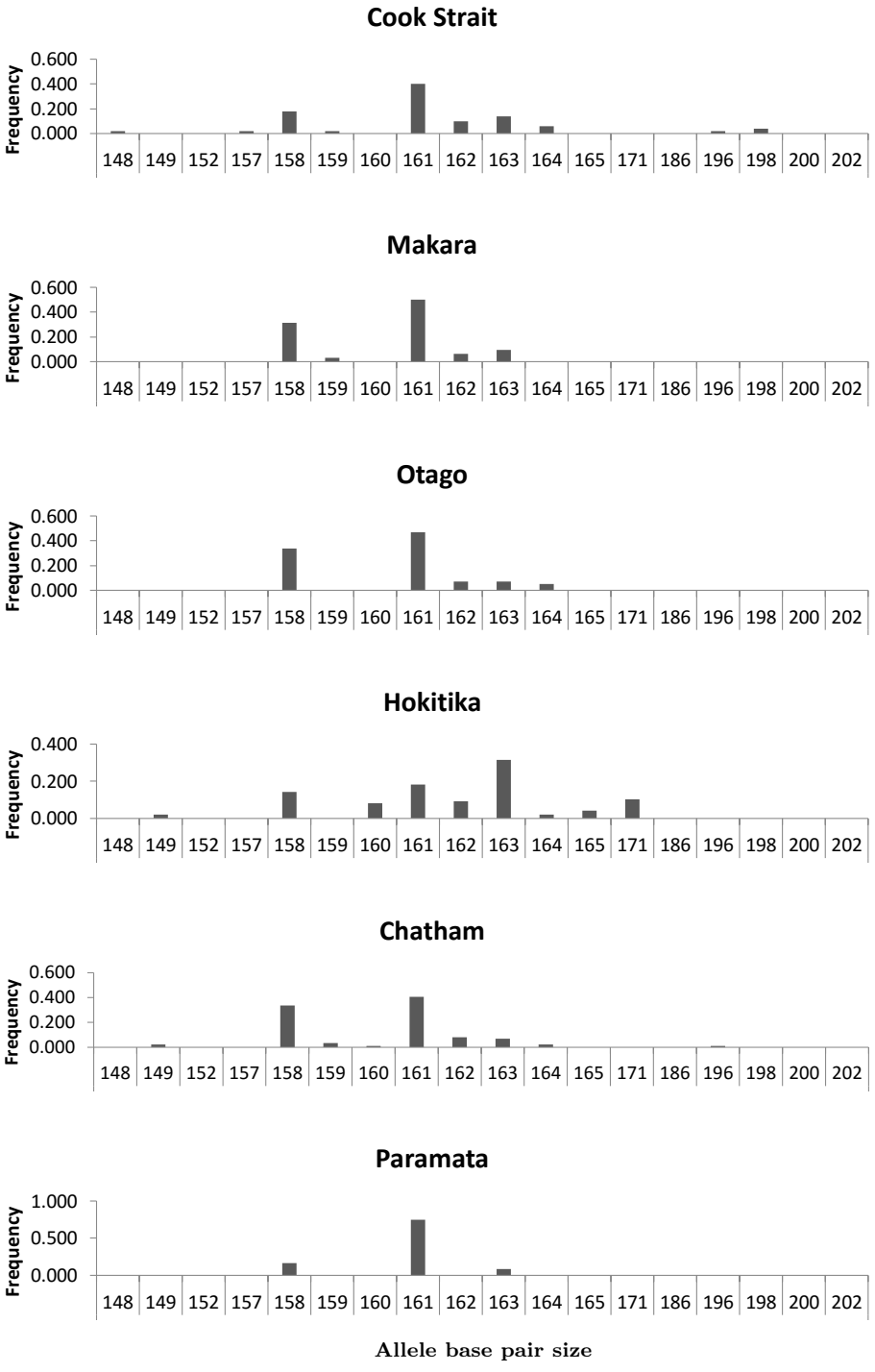



FIGURE A.2: Allelic Frequency Distribution for the microsatellite locus GG0OQ6A. Allele sizes are given in base pairs of PCR products.

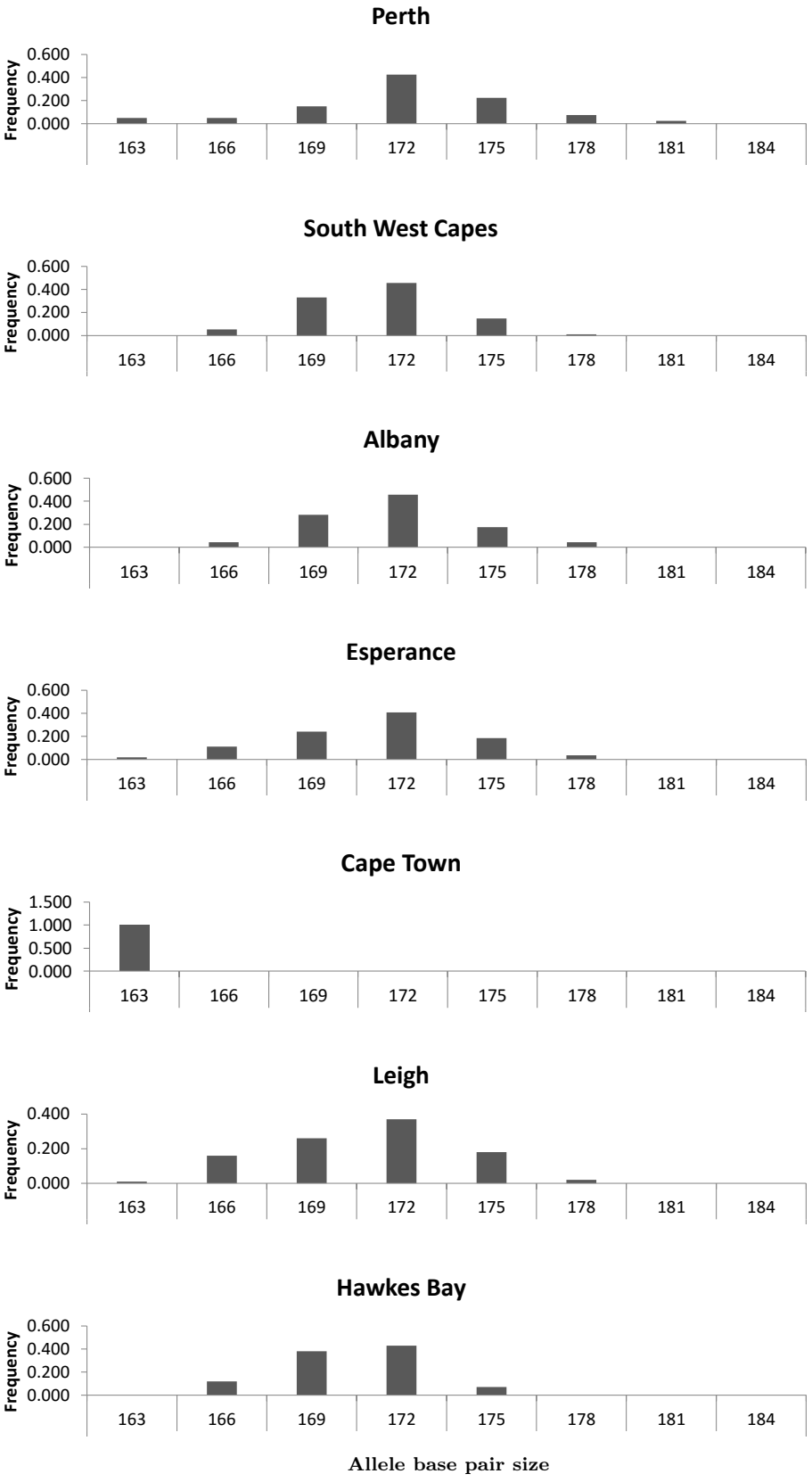



FIGURE A.3: Allelic Frequency Distribution for the microsatellite locus GH0OIK. Allele sizes are given in base pairs of PCR products.

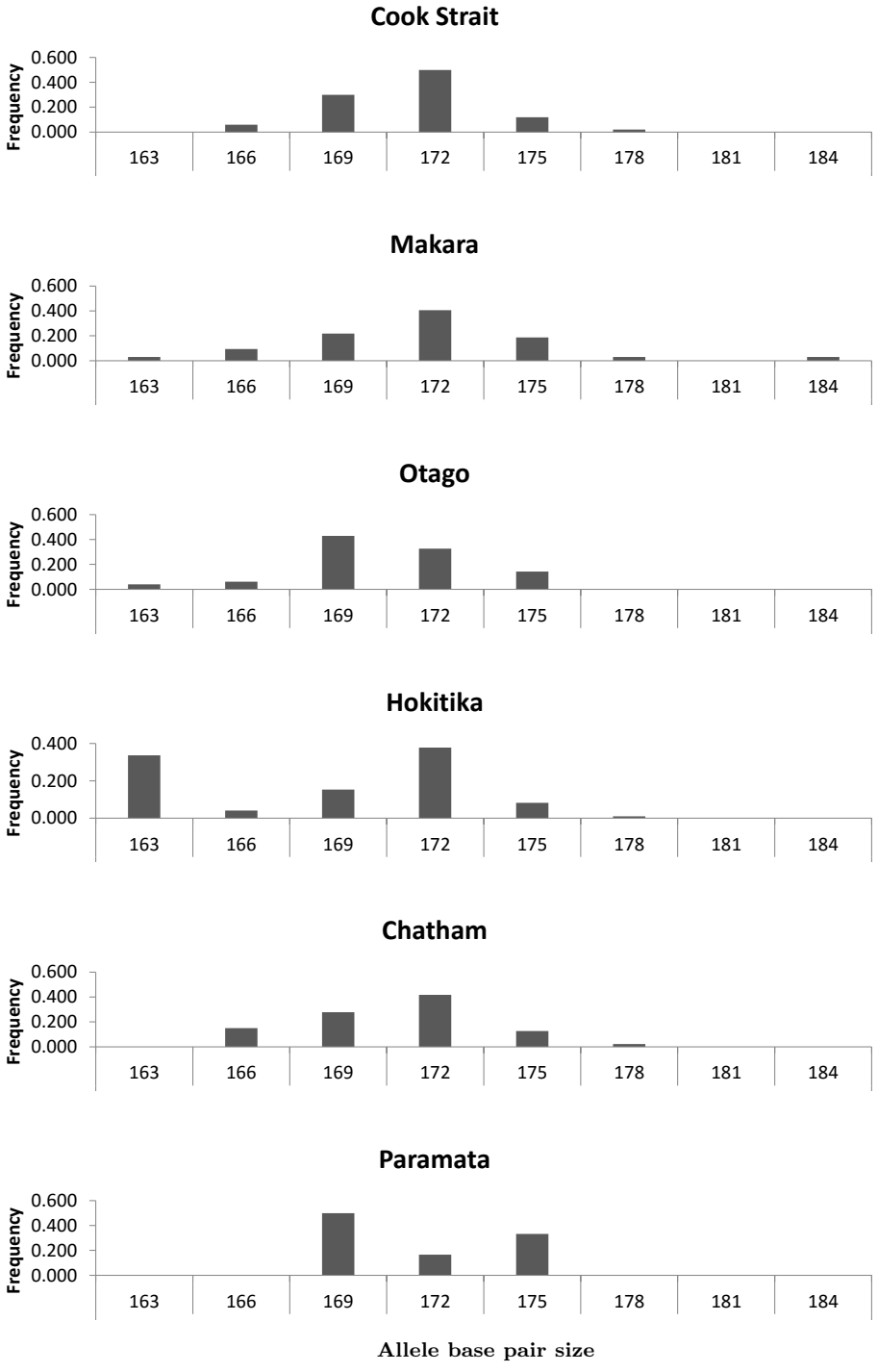



FIGURE A.4: Allelic Frequency Distribution for the microsatellite locus GH0OIK. Allele sizes are given in base pairs of PCR products.GH0OIK

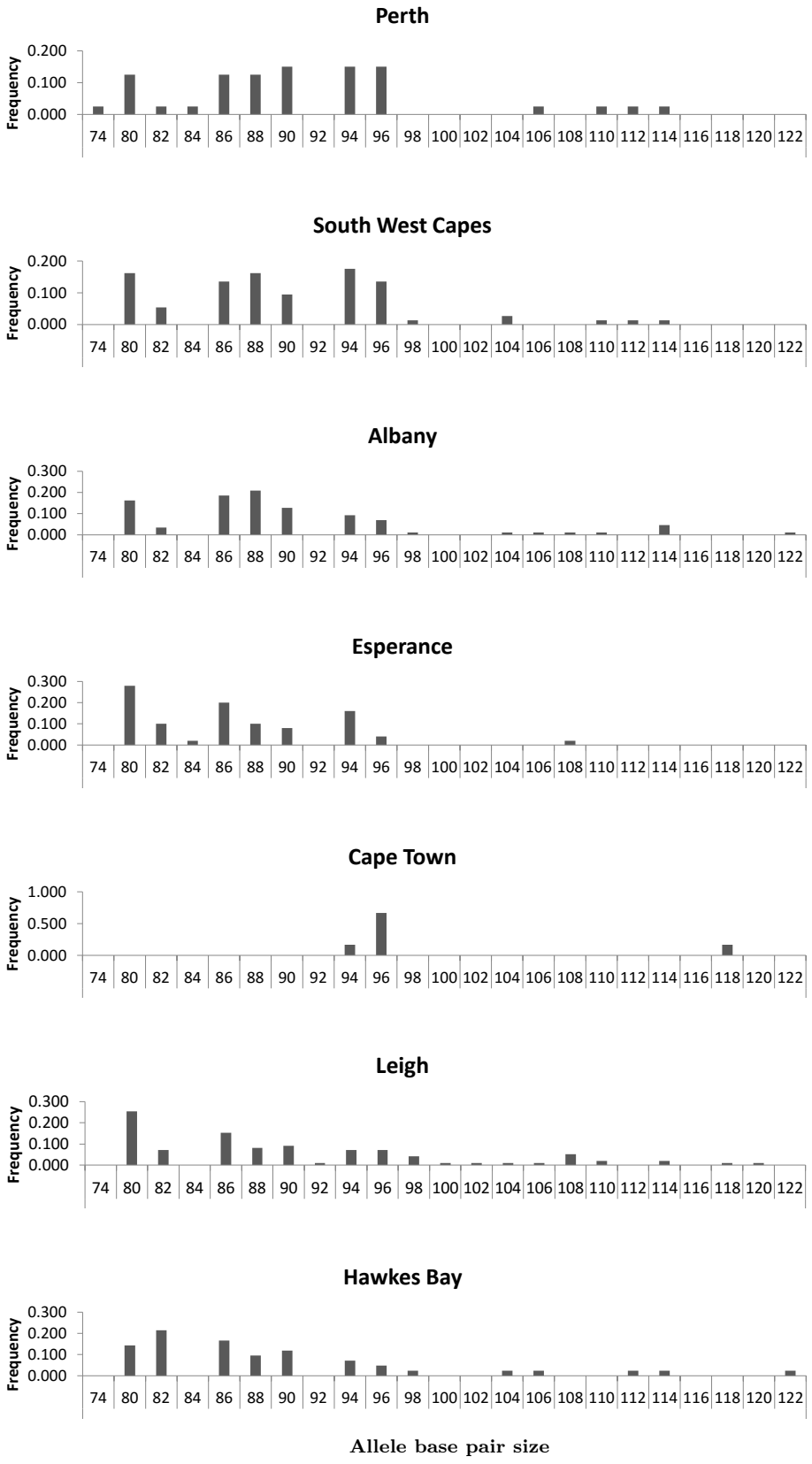



FIGURE A.5: Allelic Frequency Distribution for the microsatellite locus GJLKPX. Allele sizes are given in base pairs of PCR products.

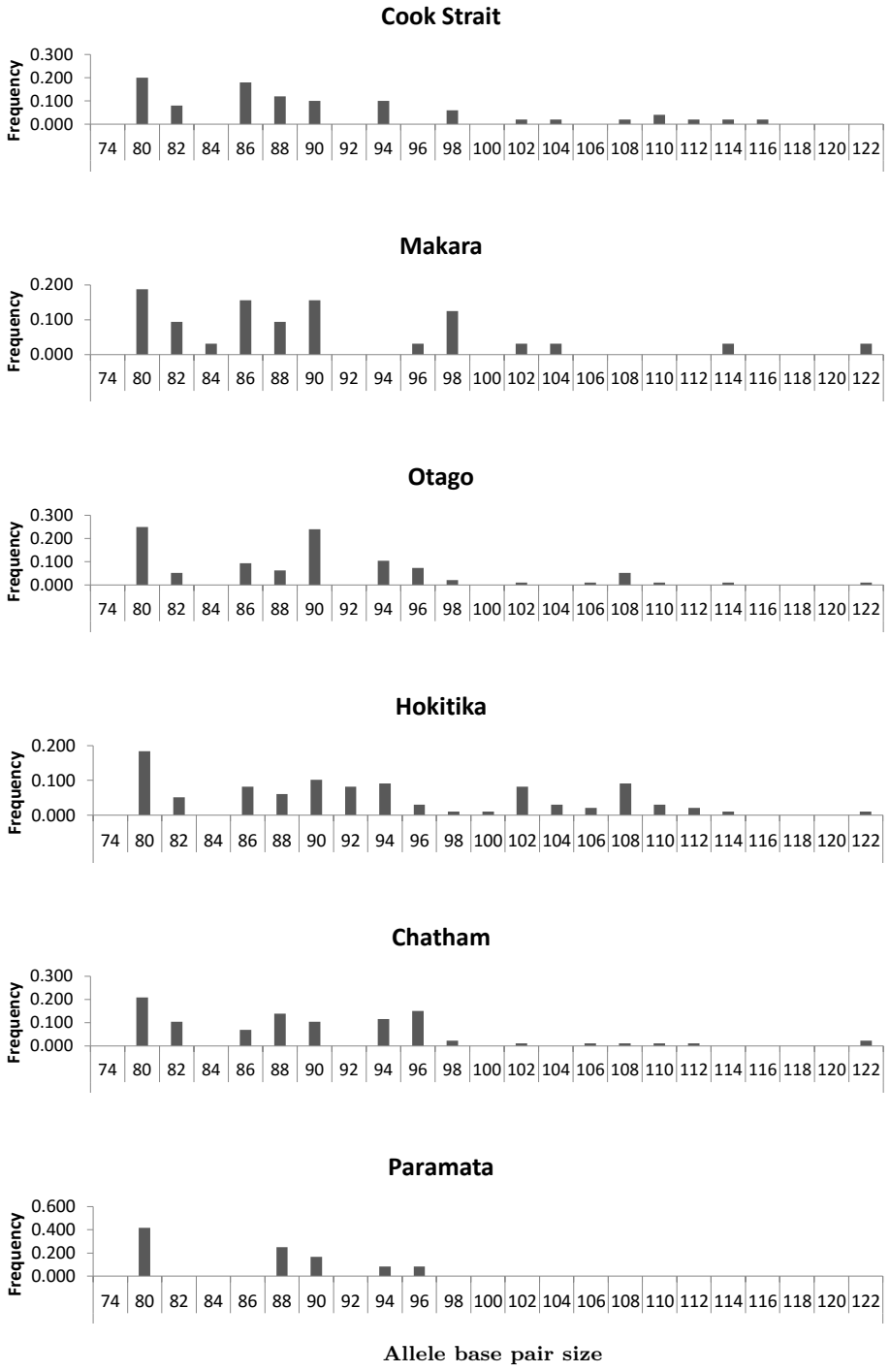



FIGURE A.6: GJLKPX

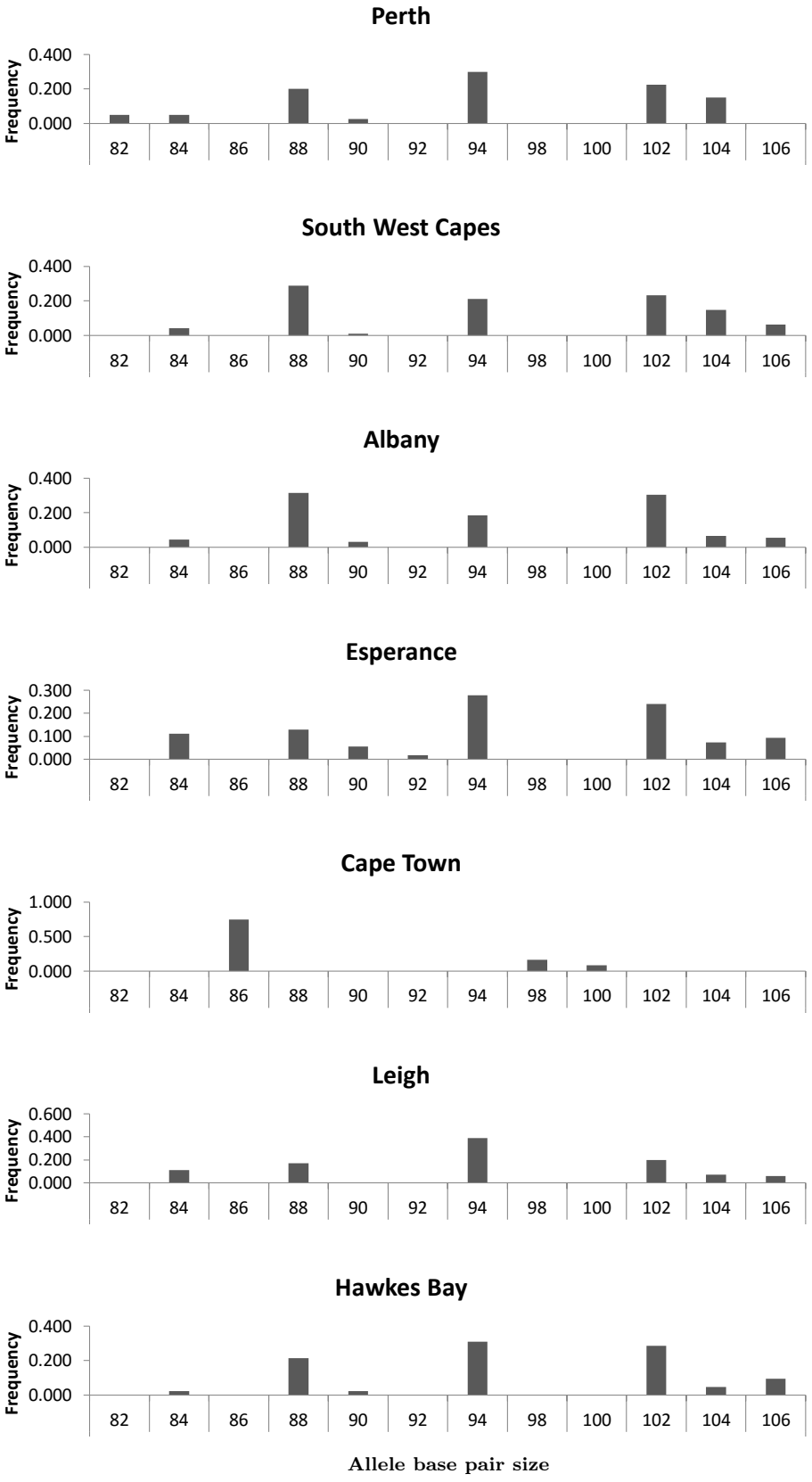



FIGURE A.7: GJSLB2

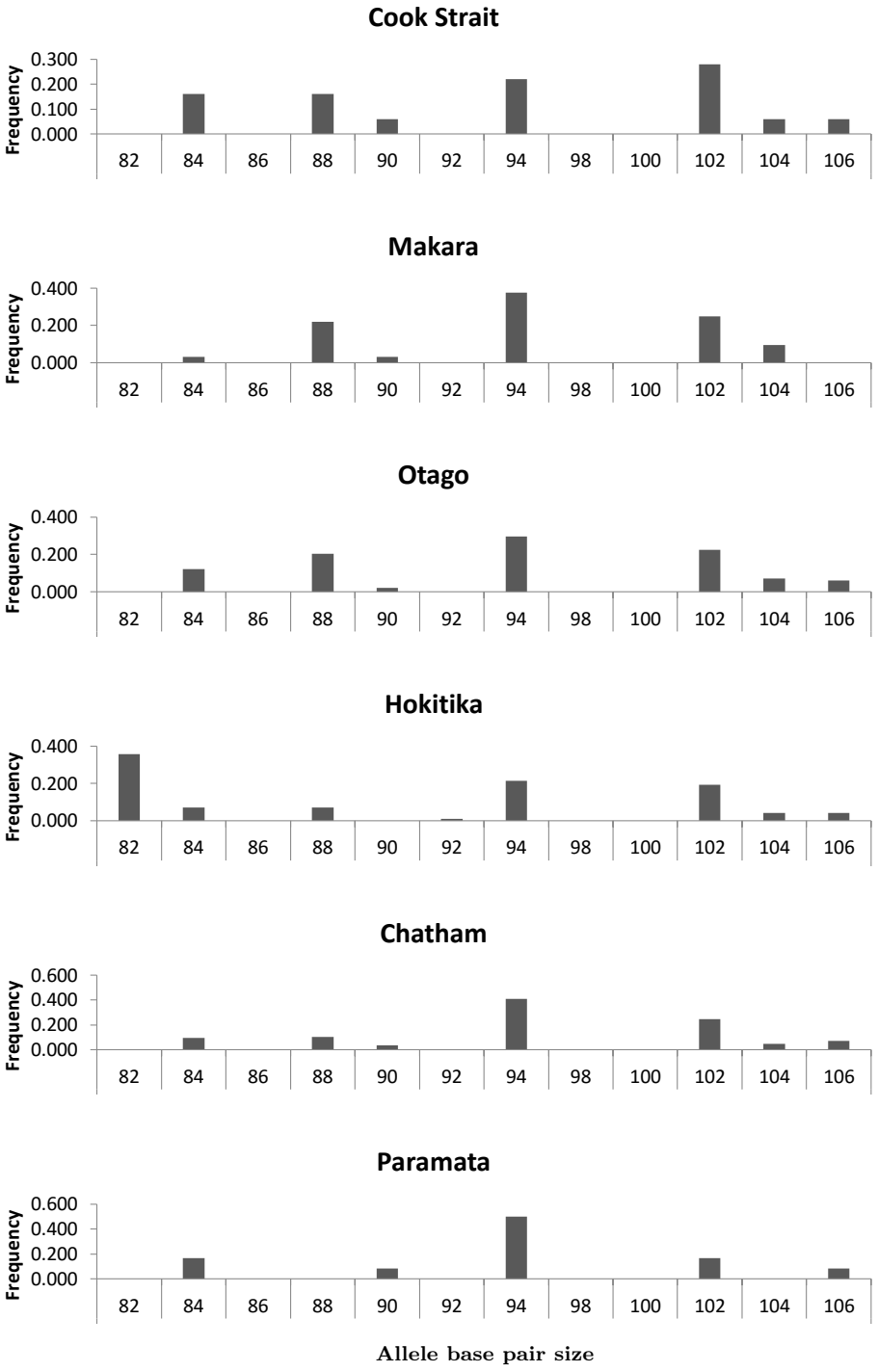



FIGURE A.8: GJSLB2

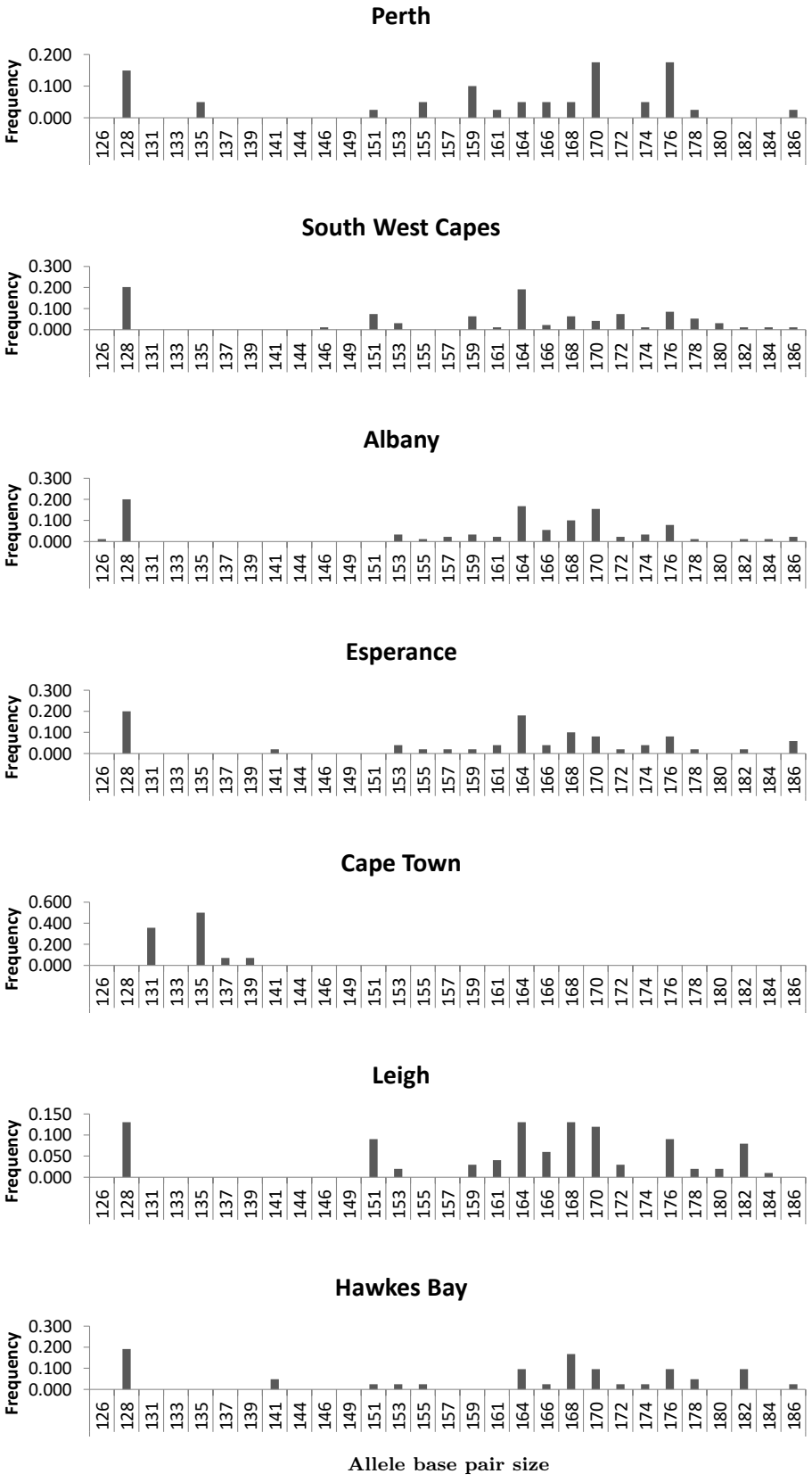



FIGURE A.9: PAM010

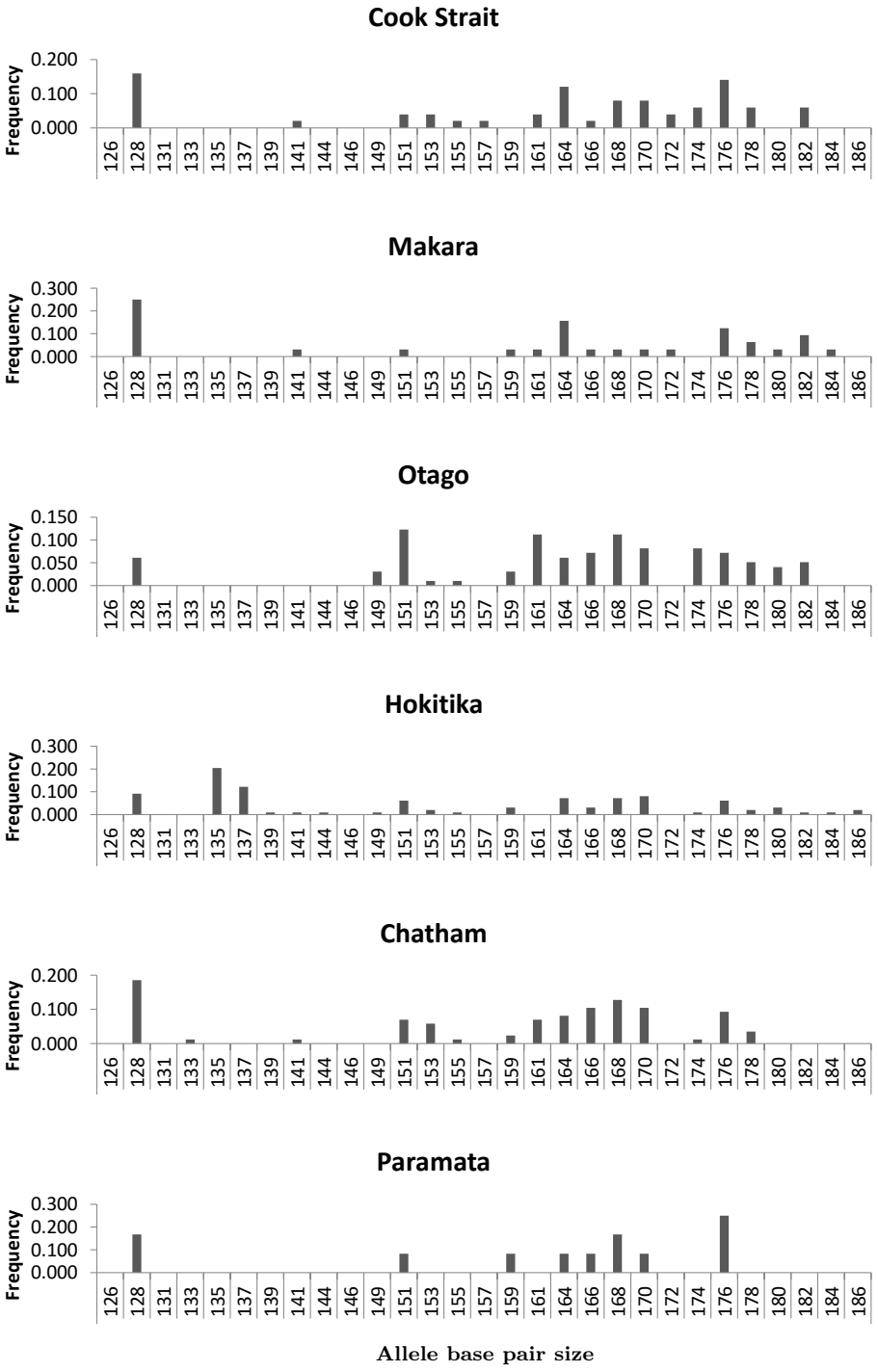



FIGURE A.10: PAM010

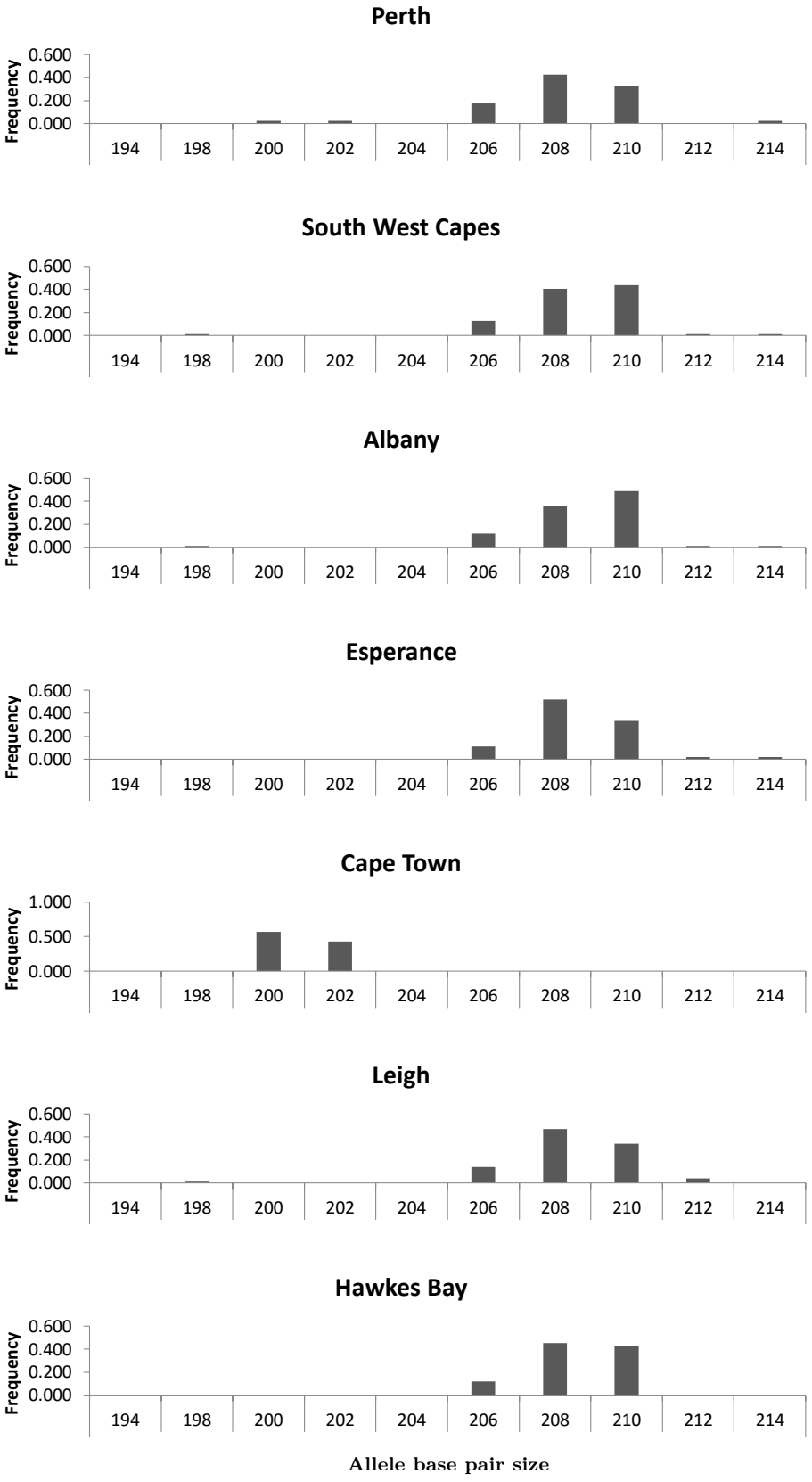



FIGURE A.11: PAM017

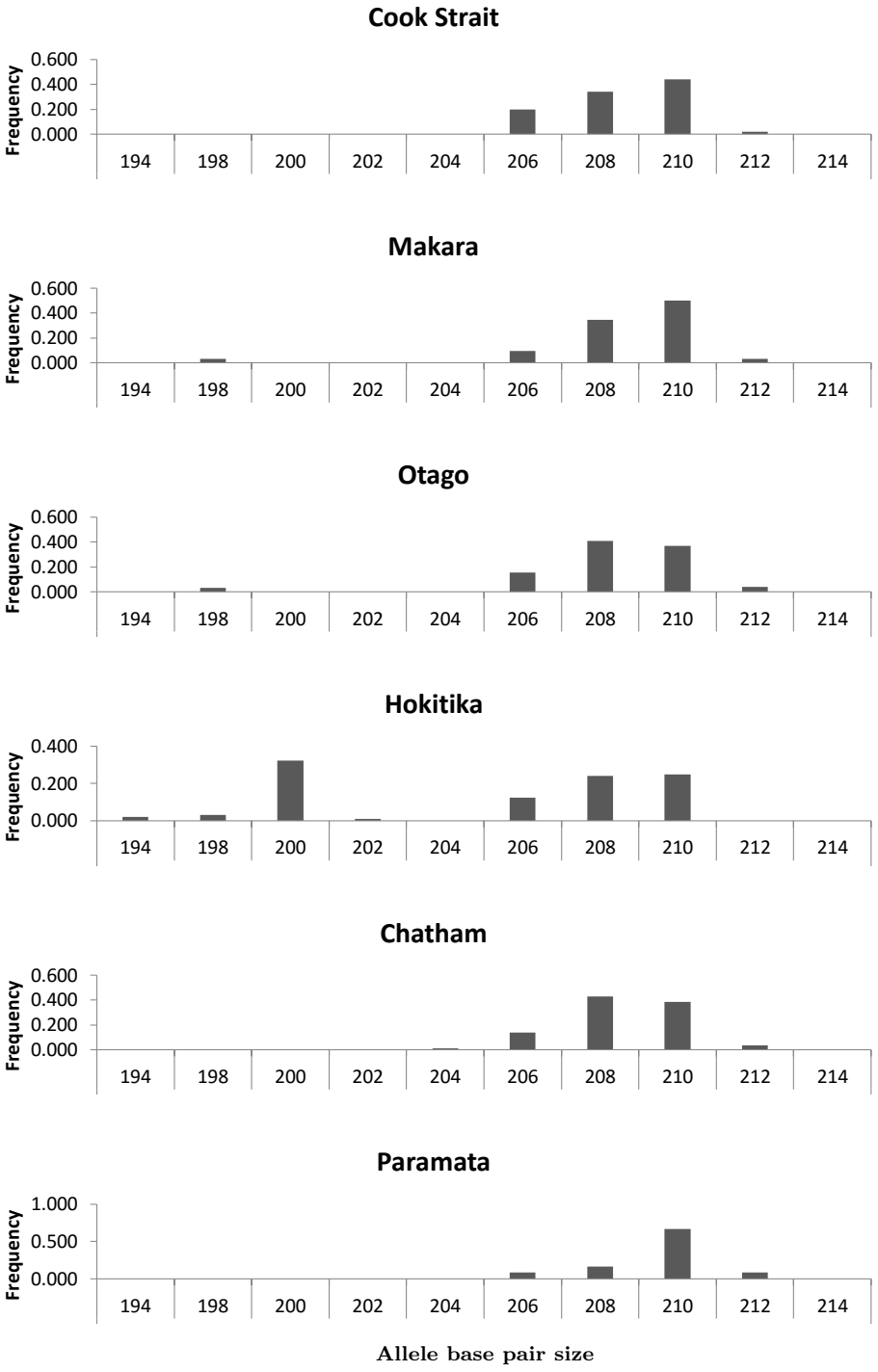



FIGURE A.12: PAM017

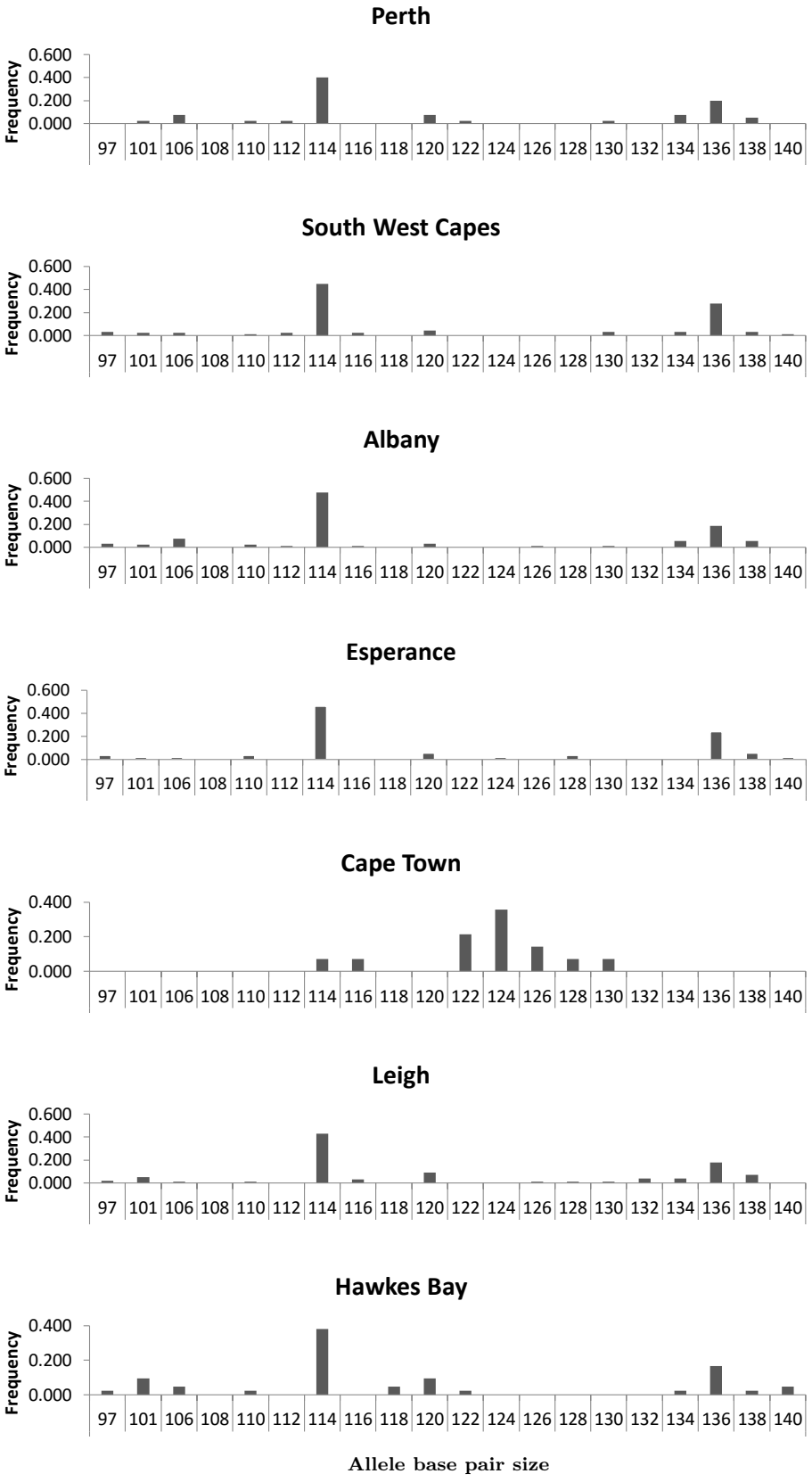



FIGURE A.13: PAM021

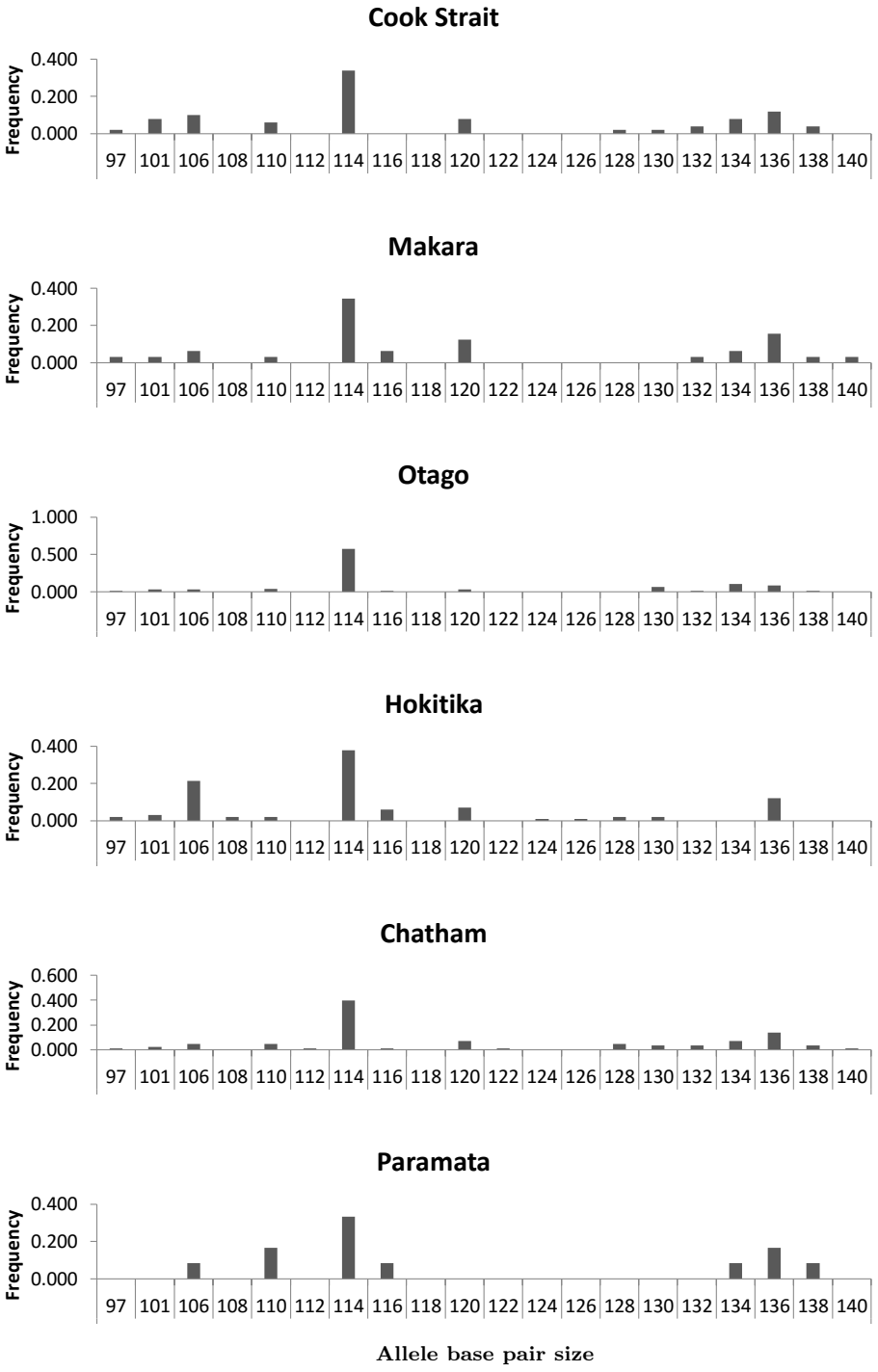



FIGURE A.14: PAM021

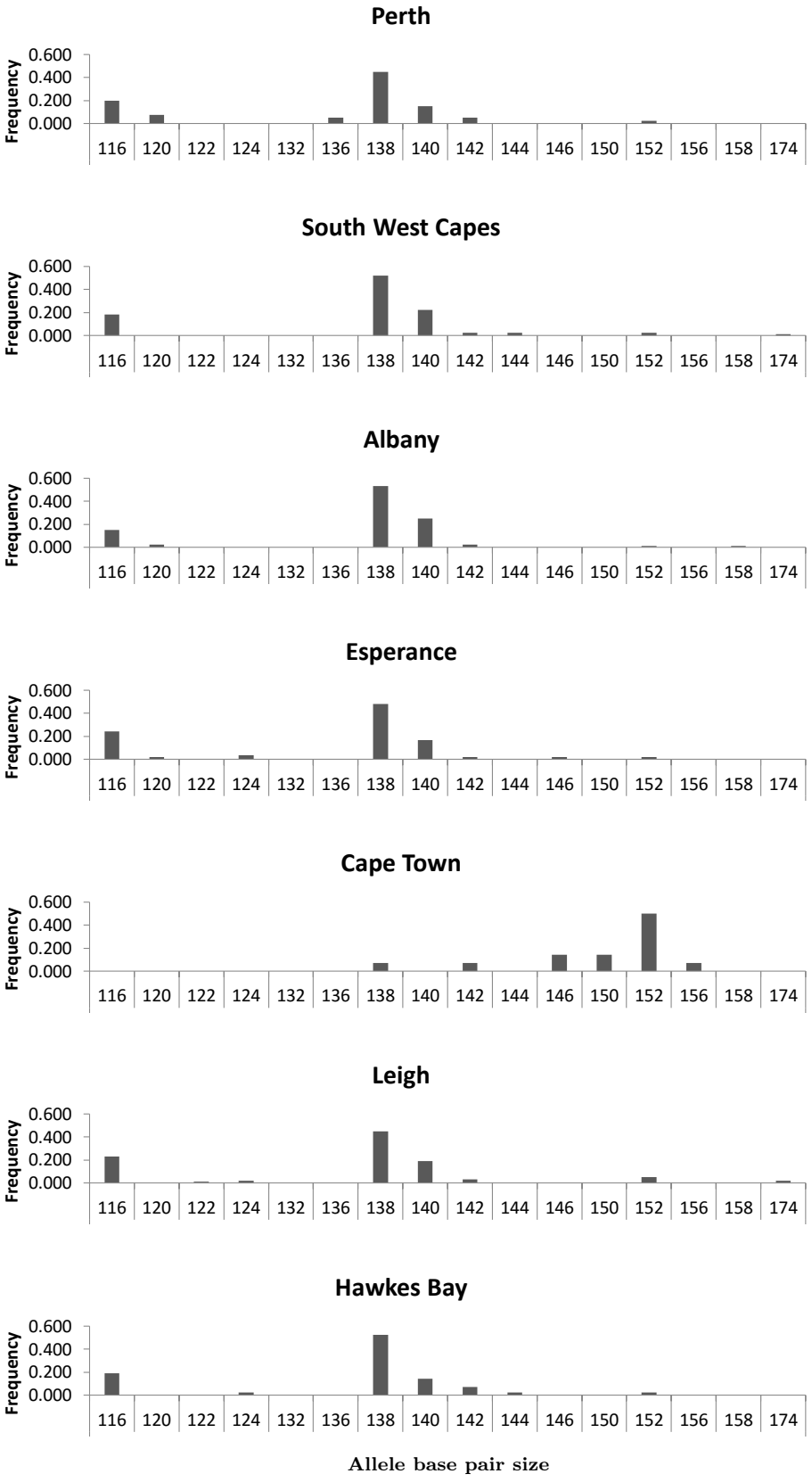



FIGURE A.15: PAM025

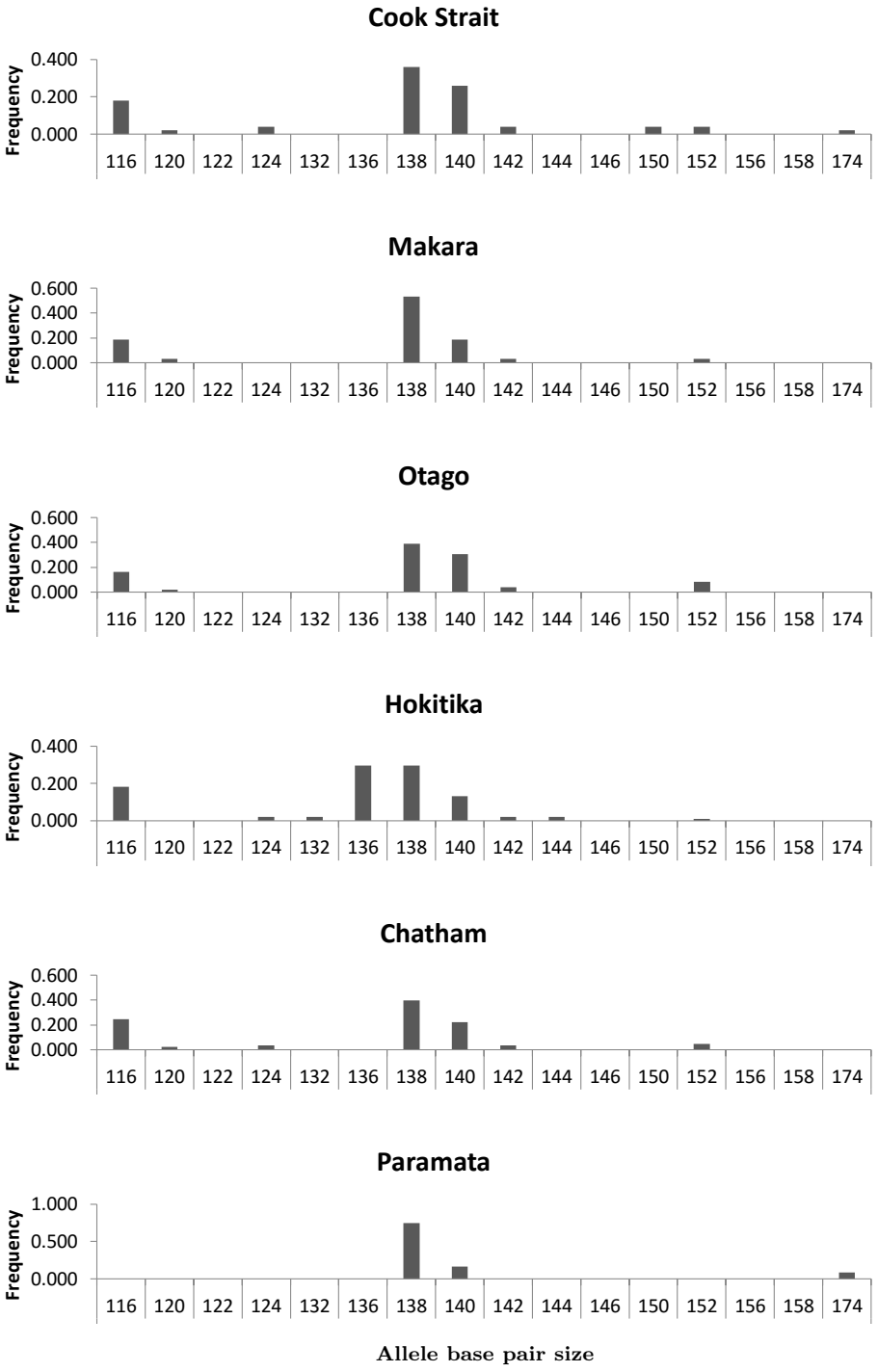



FIGURE A.16: PAM025

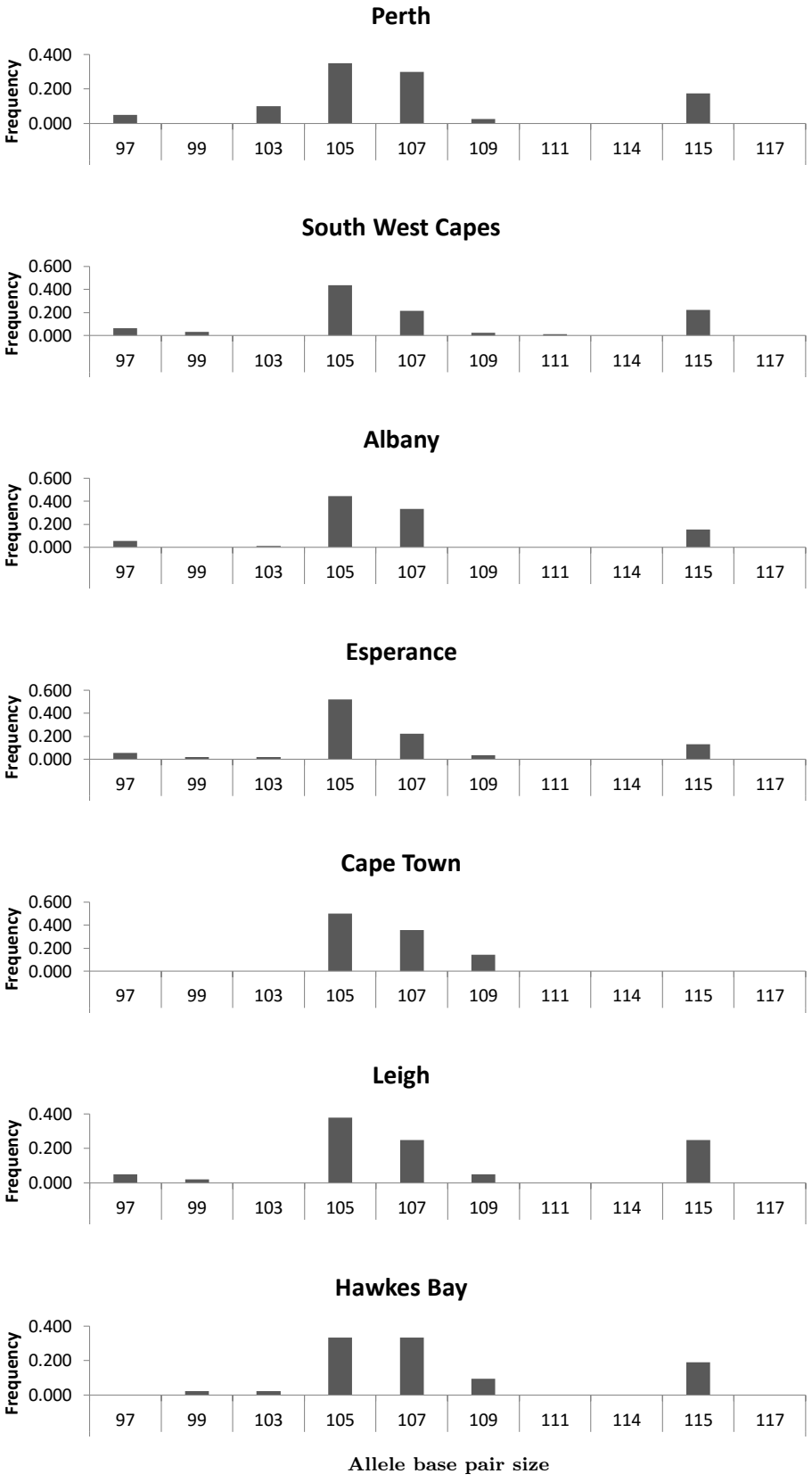



FIGURE A.17: PAM035

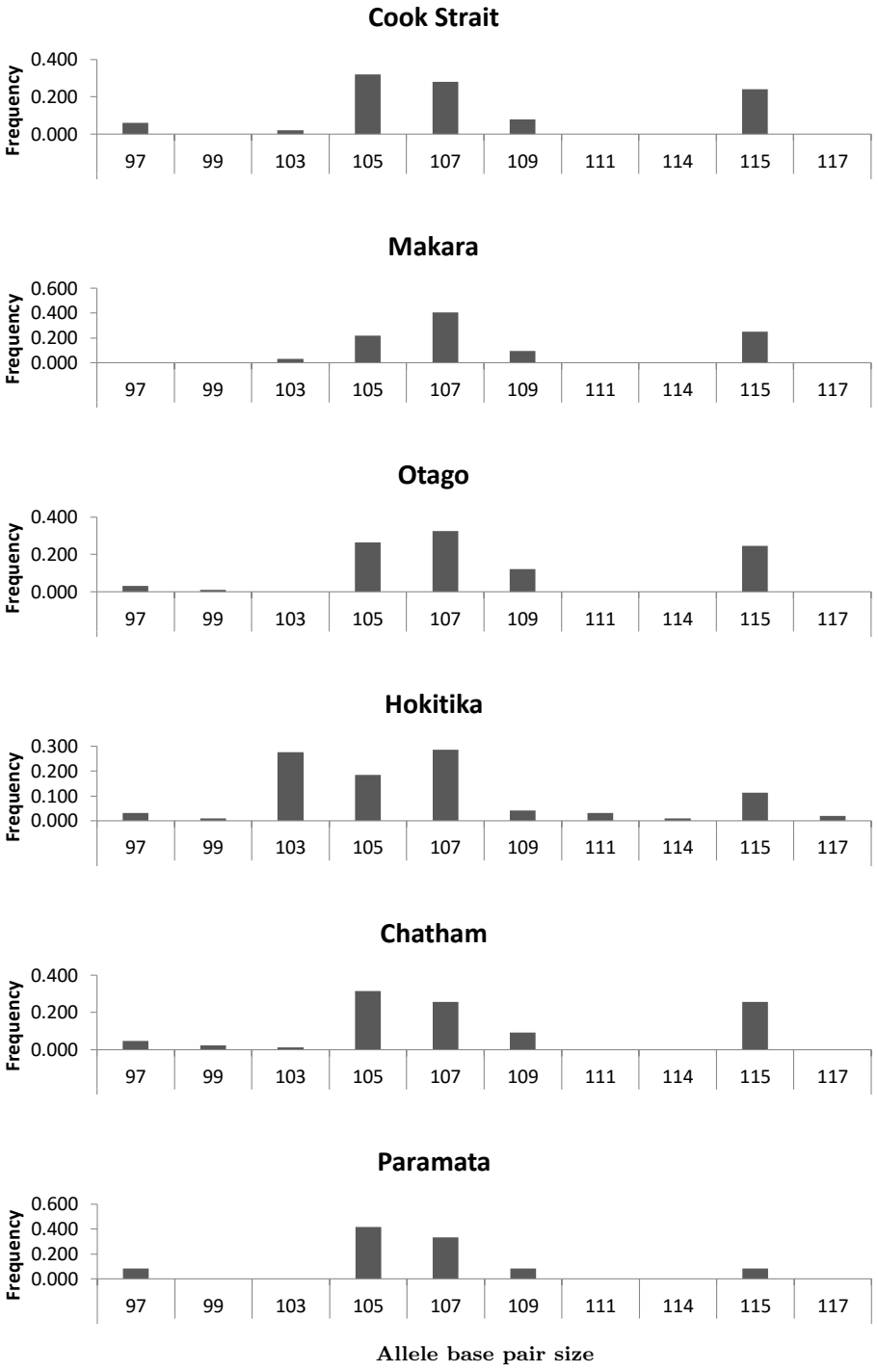



FIGURE A.18: PAM035

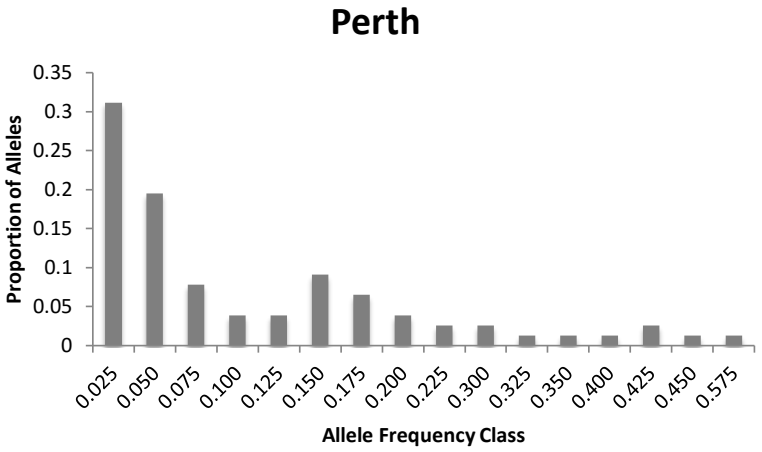



FIGURE A.19: rare allele graph

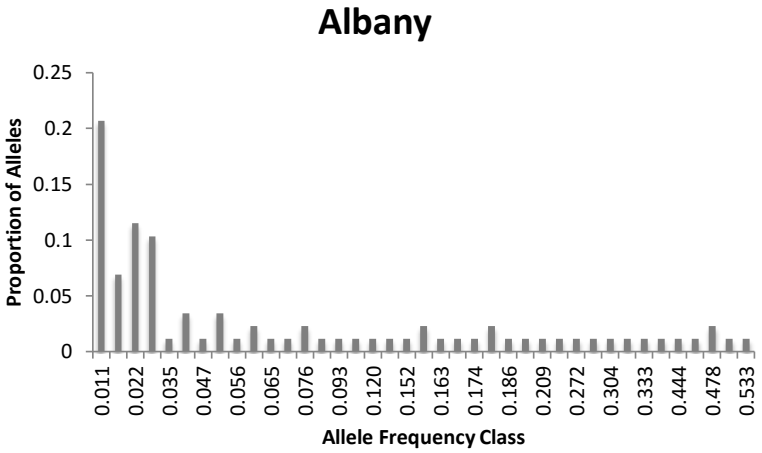



FIGURE A.20: rare allele graph

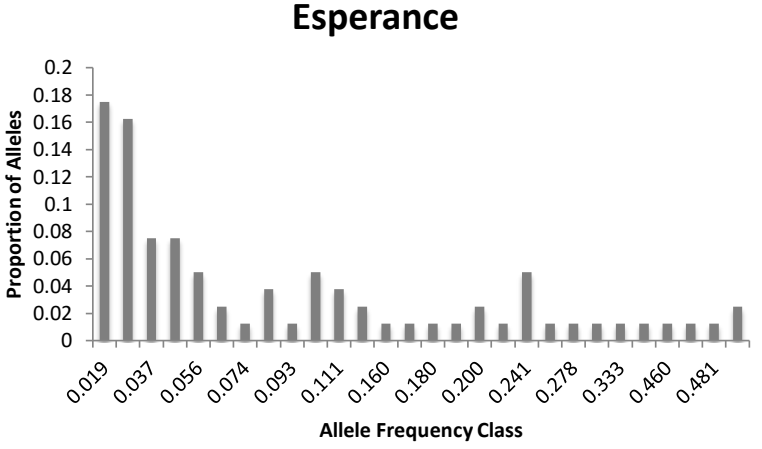



FIGURE A.21: rare allele graph

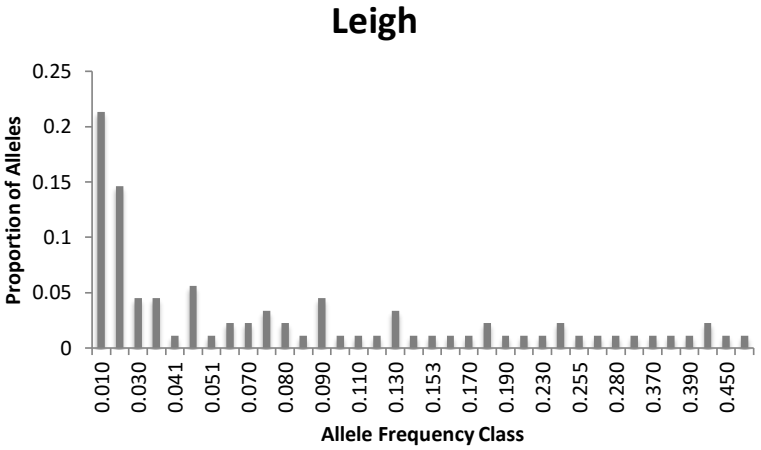



FIGURE A.22: rare allele graph

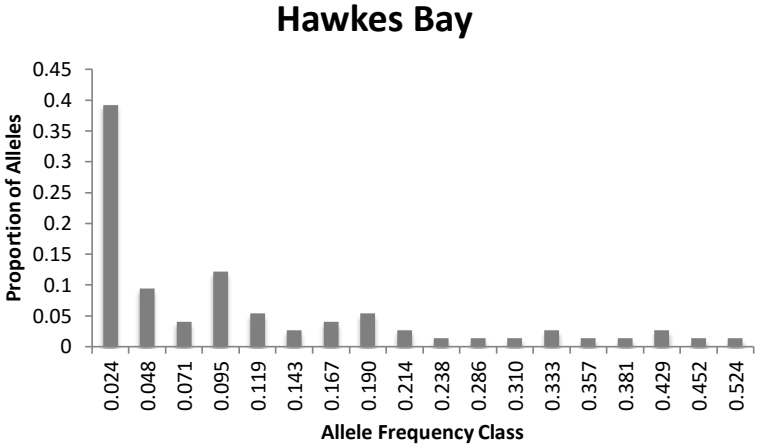



FIGURE A.23: rare allele graph

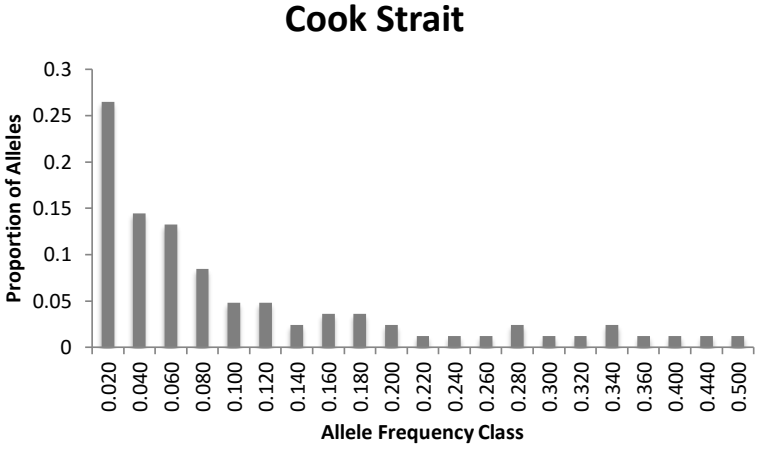



FIGURE A.24: rare allele graph

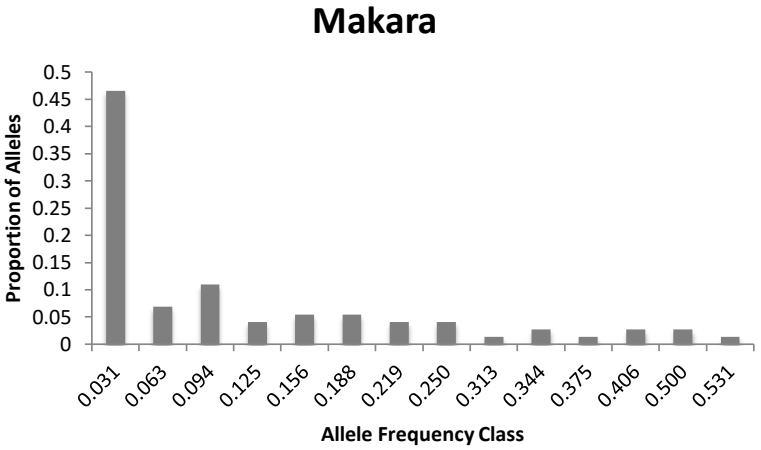



FIGURE A.25: rare allele graph

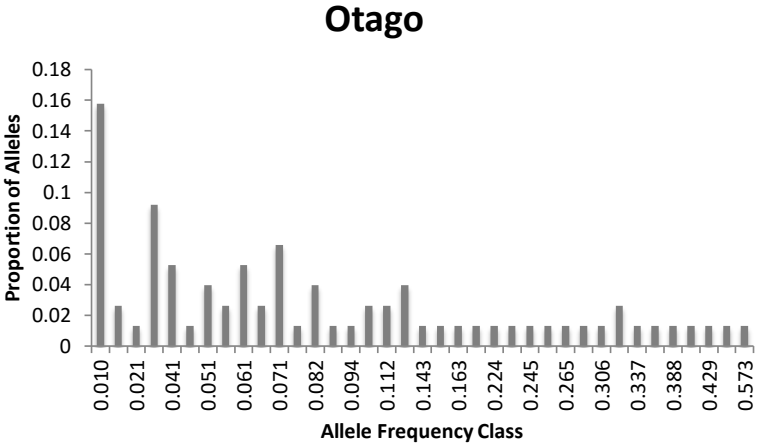



FIGURE A.26: rare allele graph

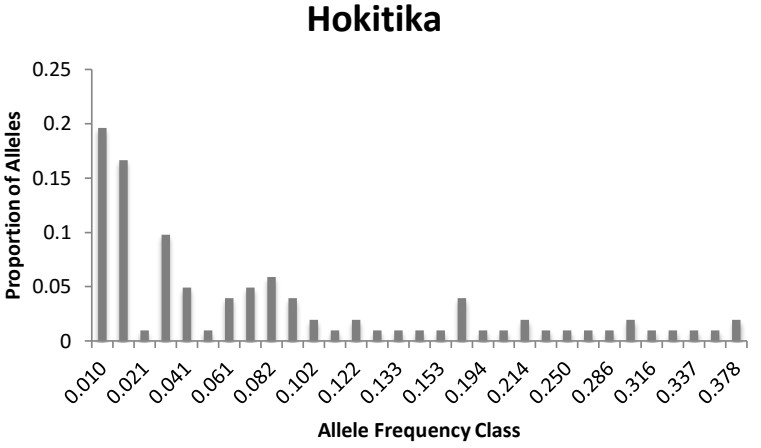



FIGURE A.27: rare allele graph

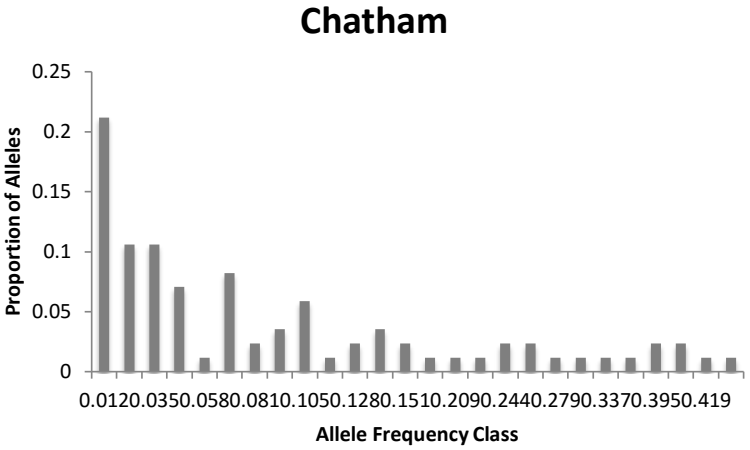



FIGURE A.28: rare allele graph

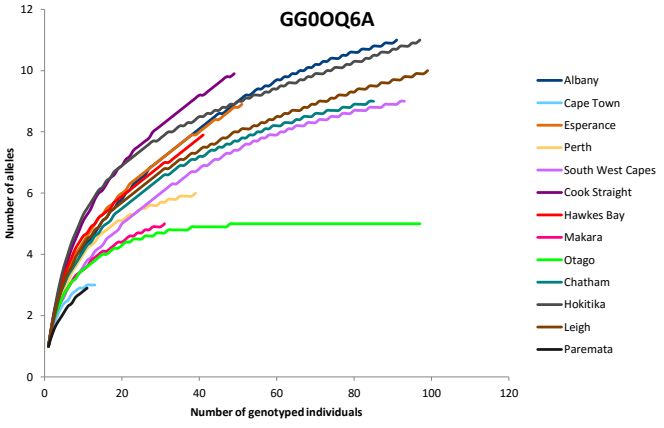



FIGURE A.29: GG0OQ6A geno

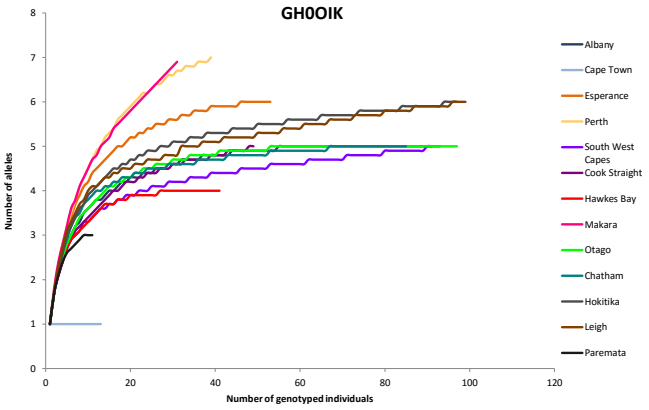



FIGURE A.30: GH0OIK geno

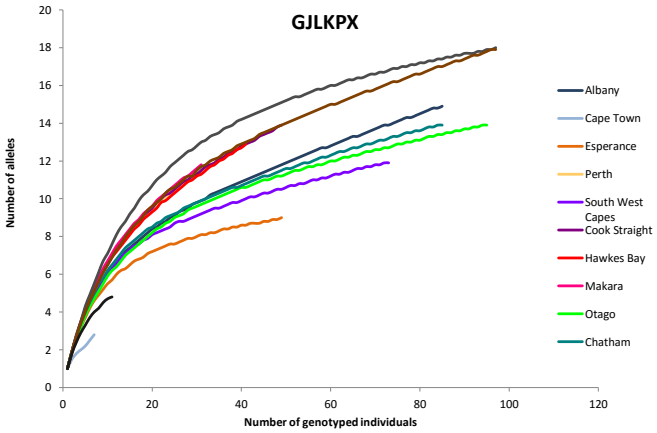



FIGURE A.31: GJLKPX geno

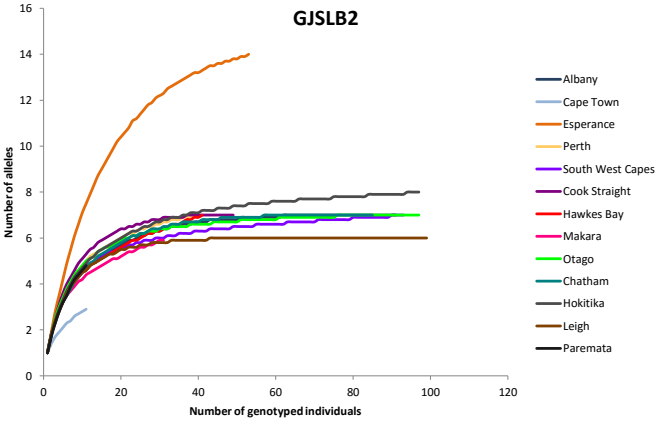



FIGURE A.32: GJSLB2 geno

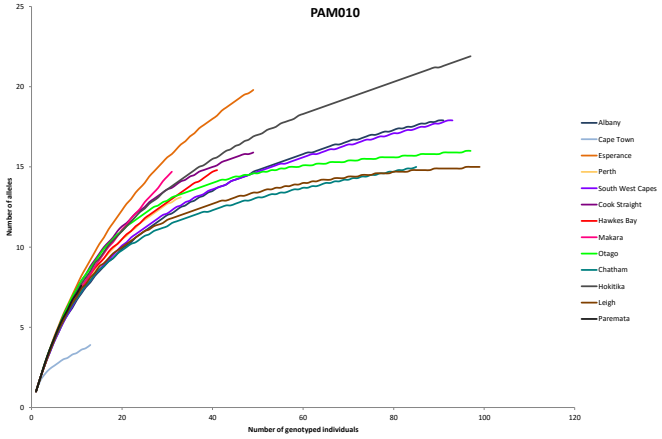



FIGURE A.33: PAM010 geno

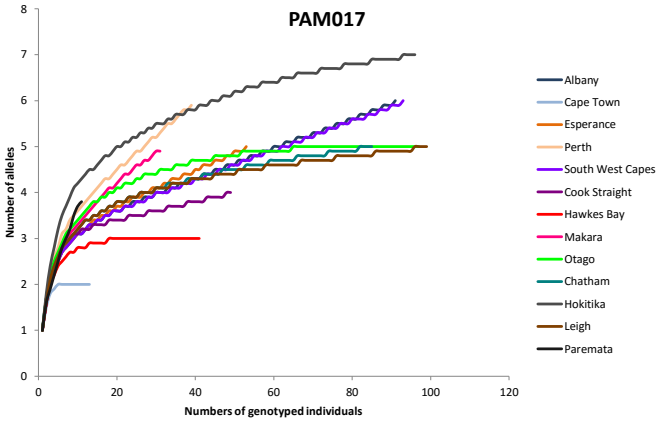



FIGURE A.34: PAM017 geno

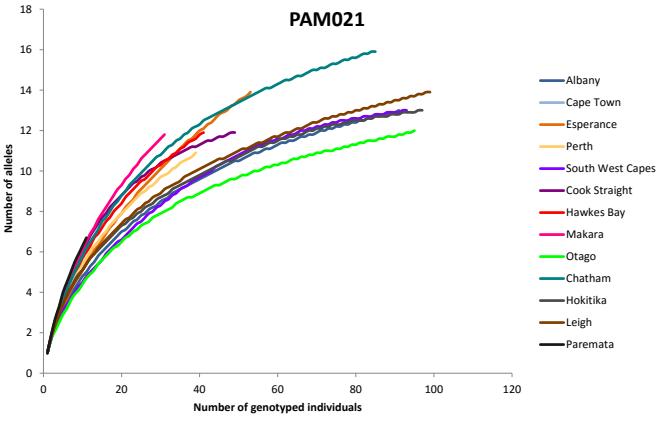



FIGURE A.35: PAM021 geno

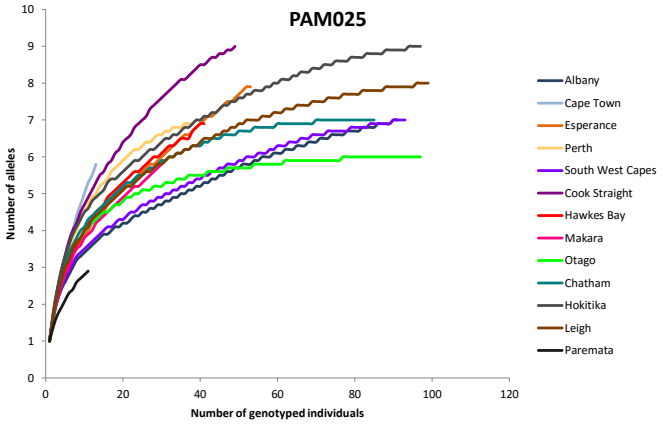



FIGURE A.36: PAM025 geno

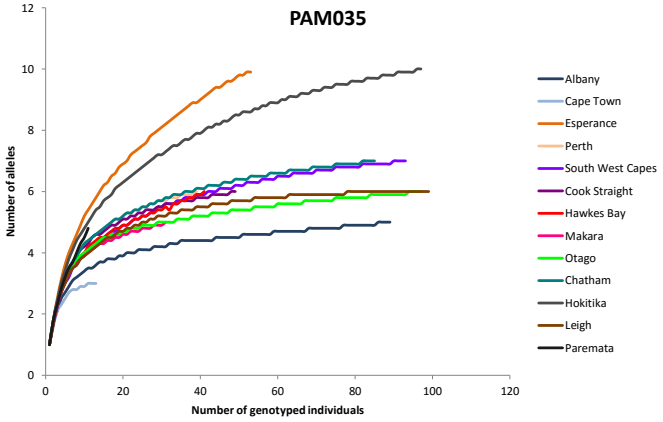



FIGURE A.37: PAM035 geno

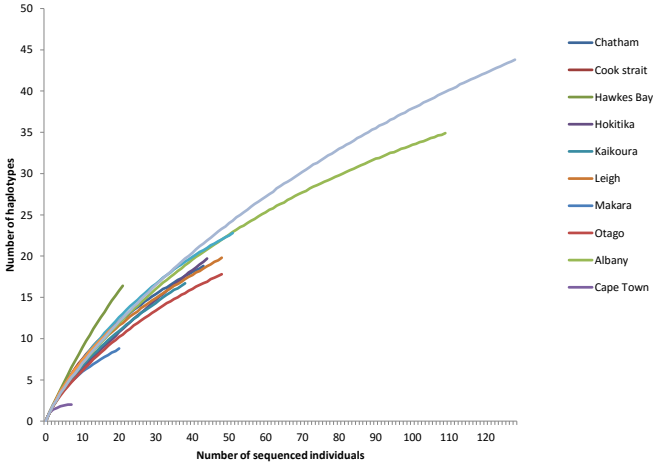



FIGURE A.38: oxygeneios seq

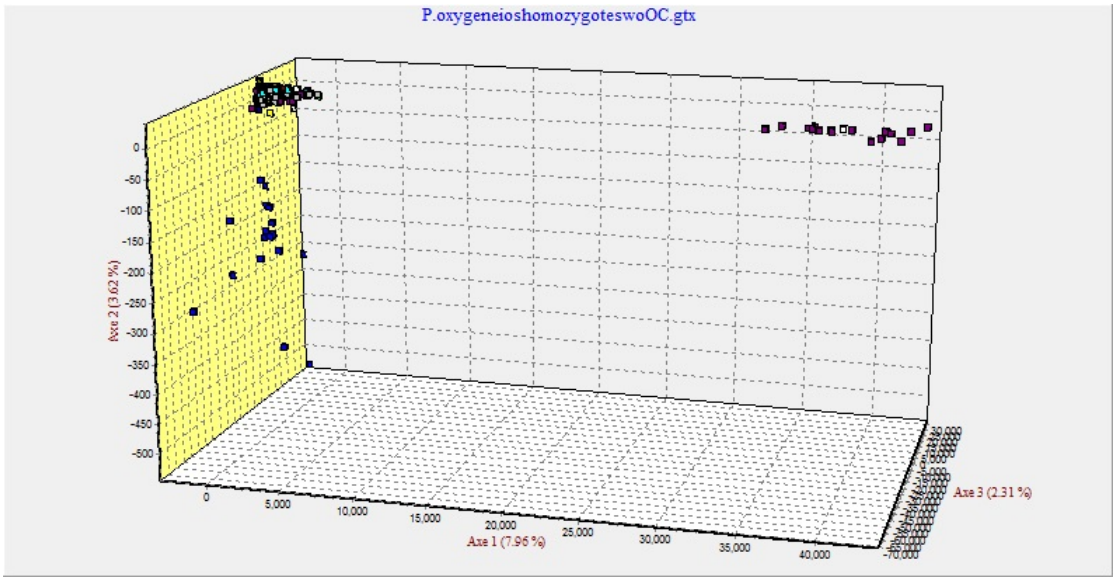



FIGURE A.39: Factorial components analysis (FCA) of *P. Oxygeneios* excluding CAP samples

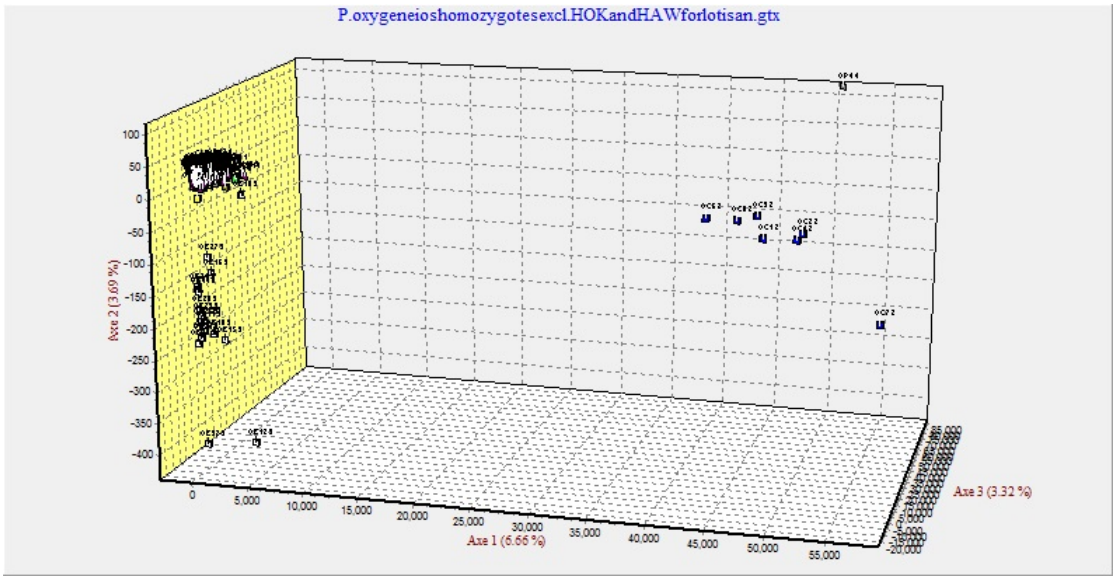



FIGURE A.40: Factorial components analysis (FCA) of *P. Oxygeneios* excluding HOK samples

TABLE A.1: G'st (Below) and Jost's D (Above) results

| | OA | OC | OE | OP | OS | OT | CS | HB | MA | OTA | CH |
|---|---|---|---|---|---|---|---|---|---|---|---|
| OA | | 0.812 | 0.13 | -0.002 | -0.006 | 0.027 | 0.004 | 0.006 | -0.002 | 0.044 | 0.031 |
| OC | 0.312 | | 0.877 | 0.751 | 0.81 | 0.751 | 0.829 | 0.819 | 0.826 | 0.807 | 0.807 |
| OE | 0.04 | 0.302 | | 0.115 | 0.126 | 0.258 | 0.105 | 0.1 | 0.115 | 0.131 | 0.095 |
| OP | -0.001 | 0.272 | 0.031 | | -0.001 | -0.003 | -0.007 | 0.002 | -0.02 | 0.036 | -0.004 |
| OS | -0.002 | 0.311 | 0.039 | 0 | | 0.029 | 0.016 | 0.013 | 0.006 | 0.045 | 0.024 |
| OT | 0.01 | 0.314 | 0.074 | -0.001 | 0.011 | | 0.025 | 0.096 | 0.06 | 0.127 | 0.042 |
| CS | 0.001 | 0.287 | 0.027 | -0.002 | 0.005 | 0.008 | | -0.023 | -0.022 | 0.019 | -0.009 |
| HB | 0.002 | 0.298 | 0.028 | 0.001 | 0.004 | 0.032 | -0.007 | | -0.028 | 0.027 | -0.009 |
| MA | -0.001 | 0.302 | 0.032 | -0.006 | 0.002 | 0.02 | -0.006 | -0.009 | | 0.022 | -0.011 |
| OTA | 0.015 | 0.302 | 0.038 | 0.011 | 0.015 | 0.043 | 0.006 | 0.009 | 0.007 | | 0.014 |
| CH | 0.01 | 0.291 | 0.026 | -0.001 | 0.008 | 0.013 | -0.002 | -0.003 | -0.003 | 0.005 | |
| HO | 0.065 | 0.213 | 0.067 | 0.037 | 0.07 | 0.05 | 0.042 | 0.05 | 0.049 | 0.06 | 0.052 |
| LE | 0.009 | 0.296 | 0.028 | -0.001 | 0.006 | 0.034 | 0 | -0.002 | -0.003 | 0.006 | -0.004 |
| PA | 0.022 | 0.338 | 0.07 | 0.02 | 0.025 | 0.065 | 0.027 | 0.029 | 0.012 | 0.028 | 0.023 |

TABLE A.2: Results from the Wilcoxon bottleneck test for *P. oxygeneios*

| | | | IAM | | | TPM | | | SMM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Population | Loci | N | H0 | He | p | H0 | He | p | H0 | He | p |
| ALB | 9 | 91 | 7 | 5.42 | 0.236 | 3 | 5.33 | 0.108 | 2 | 5.29 | 0.029 |
| CAP | 9 | 13 | 4 | 4.44 | 0.51 | 3 | 4.56 | 0.22 | 2 | 4.76 | 0.05 |
| ESP | 9 | 53 | 5 | 5.39 | 0.52 | 3 | 5.42 | 0.1 | 3 | 5.29 | 0.11 |
| PER | 9 | 40 | 8 | 5.39 | 0.07 | 4 | 5.46 | 0.25 | 2 | 5.33 | 0.03 |
| SWC | 9 | 92 | 7 | 5.35 | 0.22 | 4 | 5.35 | 0.28 | 2 | 5.36 | 0.03 |
| COO | 9 | 50 | 9 | 5.36 | 0.01 | 6 | 5.34 | 0.46 | 4 | 5.32 | 0.28 |
| HAW | 9 | 42 | 7 | 5.34 | 0.22 | 6 | 5.41 | 0.48 | 2 | 5.28 | 0.03 |
| MAK | 9 | 32 | 6 | 5.4 | 0.48 | 3 | 5.43 | 0.1 | 1 | 5.42 | 0.004 |
| OTA | 9 | 98 | 8 | 5.35 | 0.07 | 7 | 5.3 | 0.21 | 6 | 5.33 | 0.46 |
| CHA | 9 | 86 | 8 | 5.37 | 0.07 | 7 | 5.38 | 0.23 | 2 | 5.33 | 0.03 |
| HOK | 9 | 98 | 8 | 5.4 | 0.07 | 7 | 5.34 | 0.3 | 0 | 5.35 | 0.0003 |
| LEI | 9 | 100 | 7 | 5.36 | 0.22 | 5 | 5.31 | 0.54 | 4 | 5.32 | 0.29 |

IAM=infinate alleles model, TPM=two phase model, SMM=stepwise mutation model, N=sample size, H0=observed heterozygosity, He=expected heterozygosity, and p=p-value

# Appendix B

# Additional information for *P. americanus*

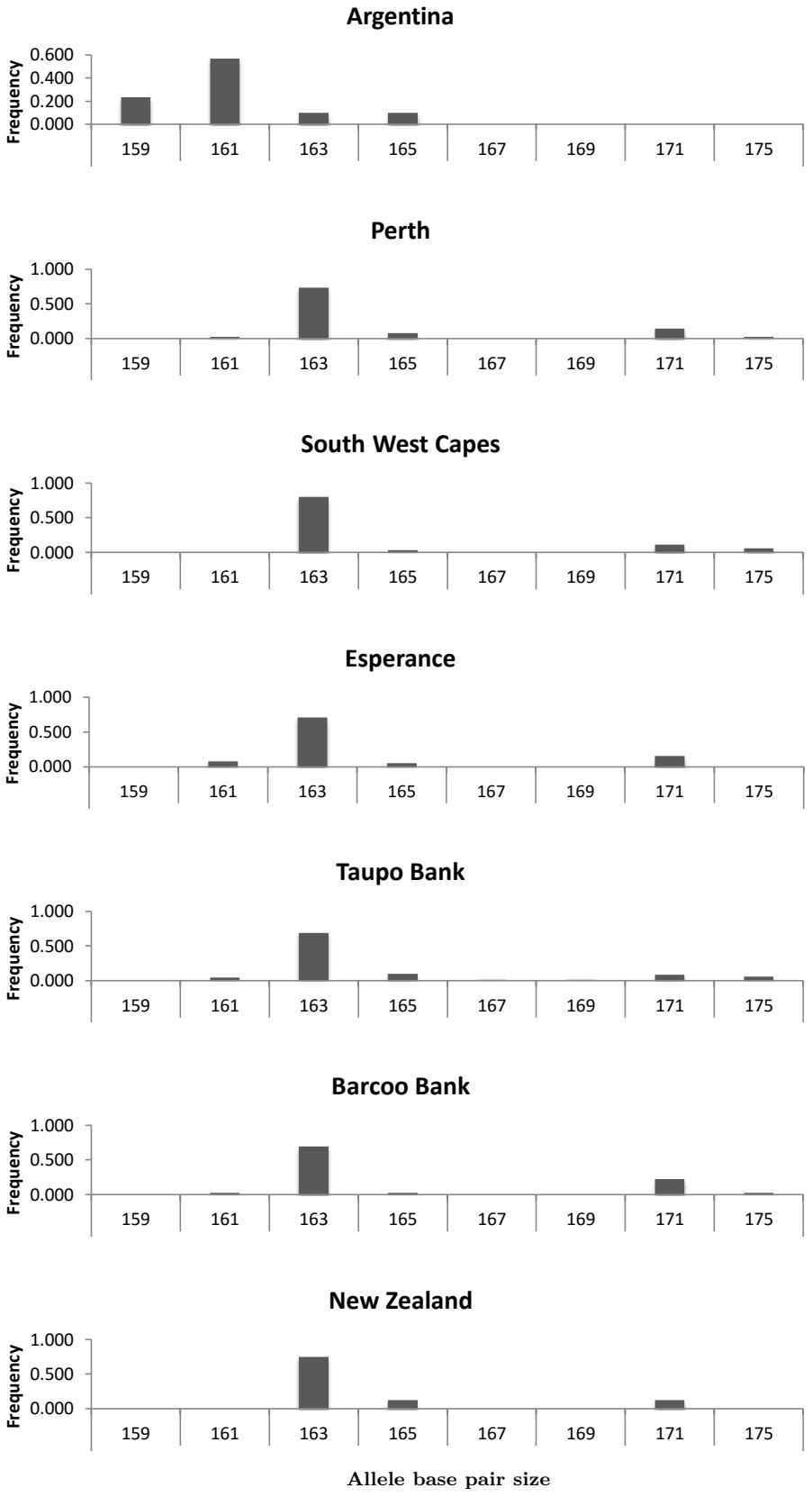



FIGURE B.1: Allelic Frequency Distribution for the microsatellite locus GGOQ6A. Allele sizes are given in base pairs of PCR products.

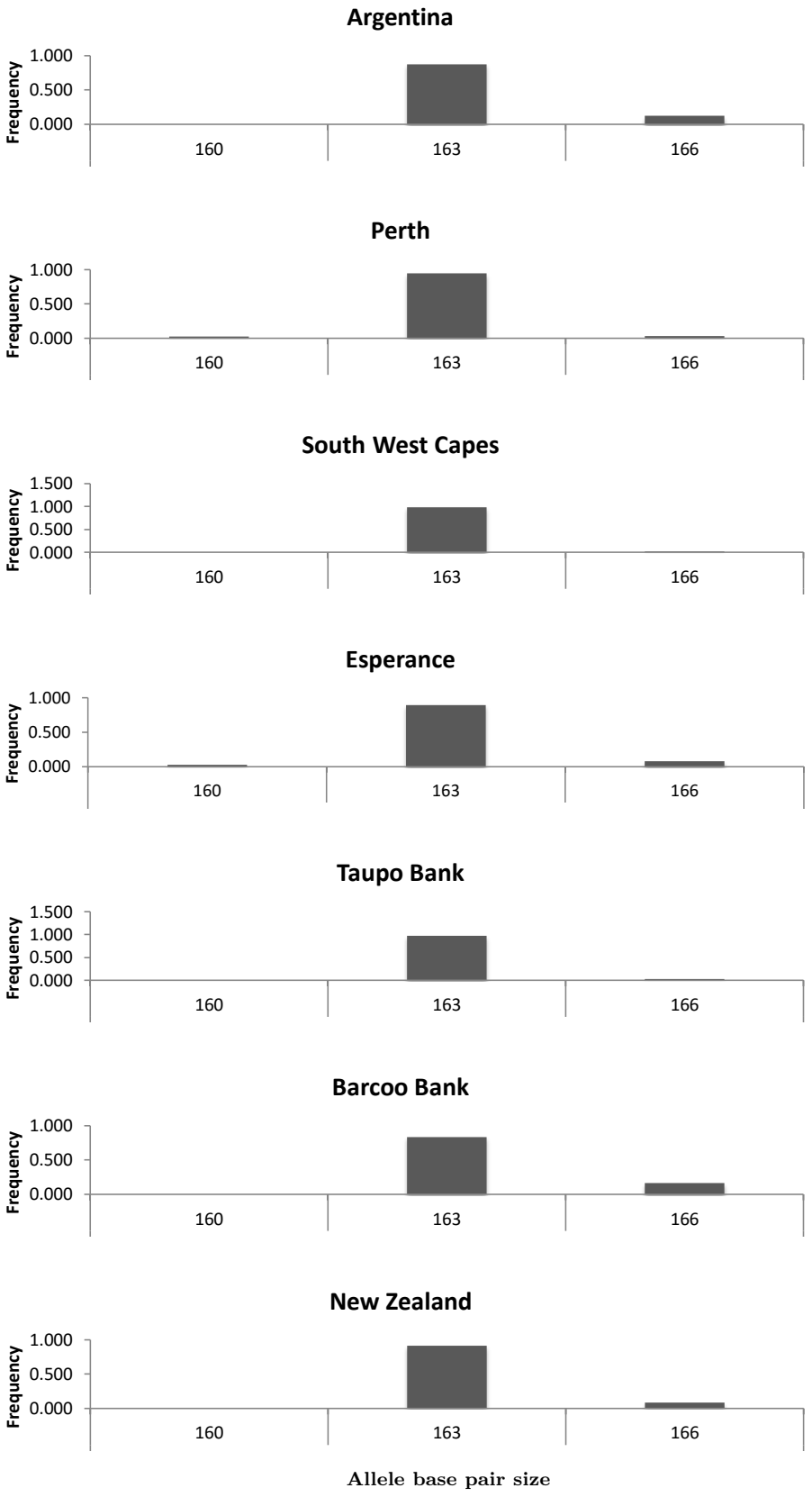



FIGURE B.2: Allelic Frequency Distribution for the microsatellite locus GH0OIK. Allele sizes are given in base pairs of PCR products.

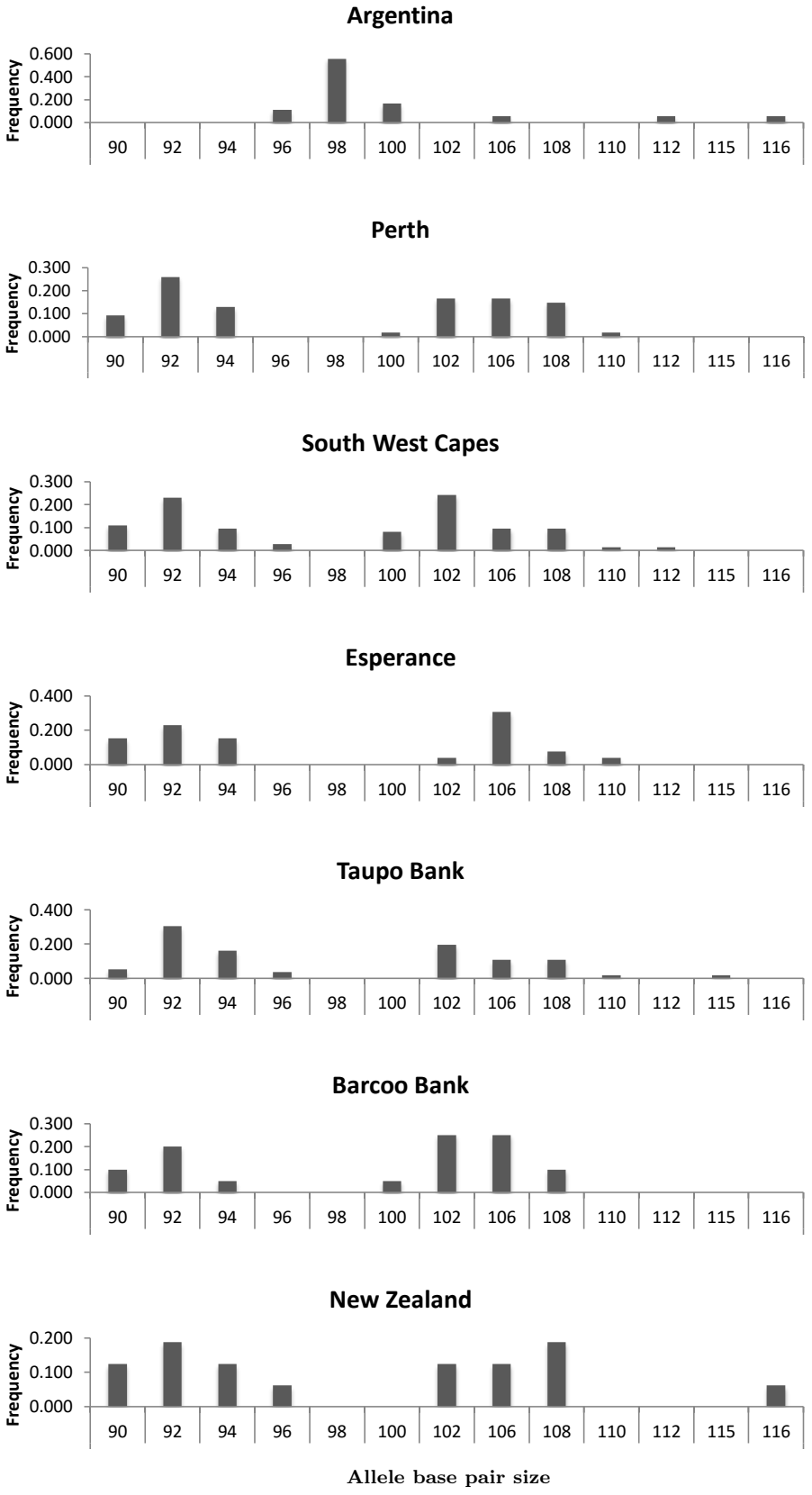



FIGURE B.3: Allelic Frequency Distribution for the microsatellite locus GJLKPX. Allele sizes are given in base pairs of PCR products7.

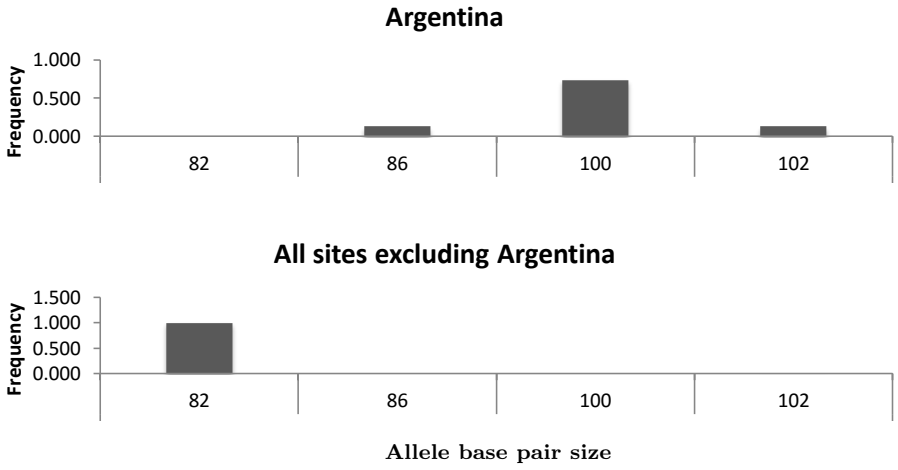



FIGURE B.4: Allelic Frequency Distribution for the microsatellite locus GJSLB2. Allele sizes are given in base pairs of PCR products.

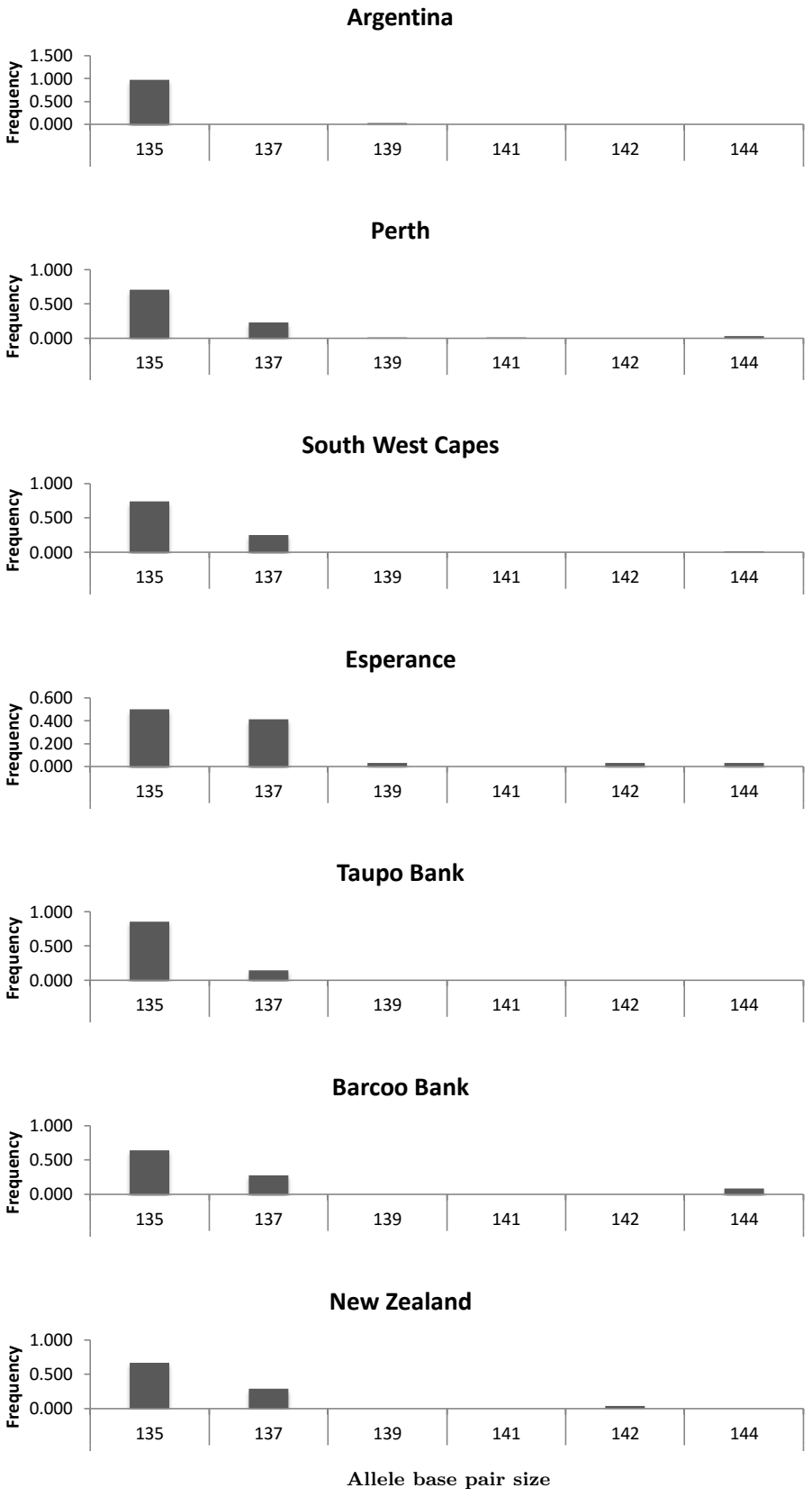



FIGURE B.5: Allelic Frequency Distribution for the microsatellite locus PAM010. Allele sizes are given in base pairs of PCR products.

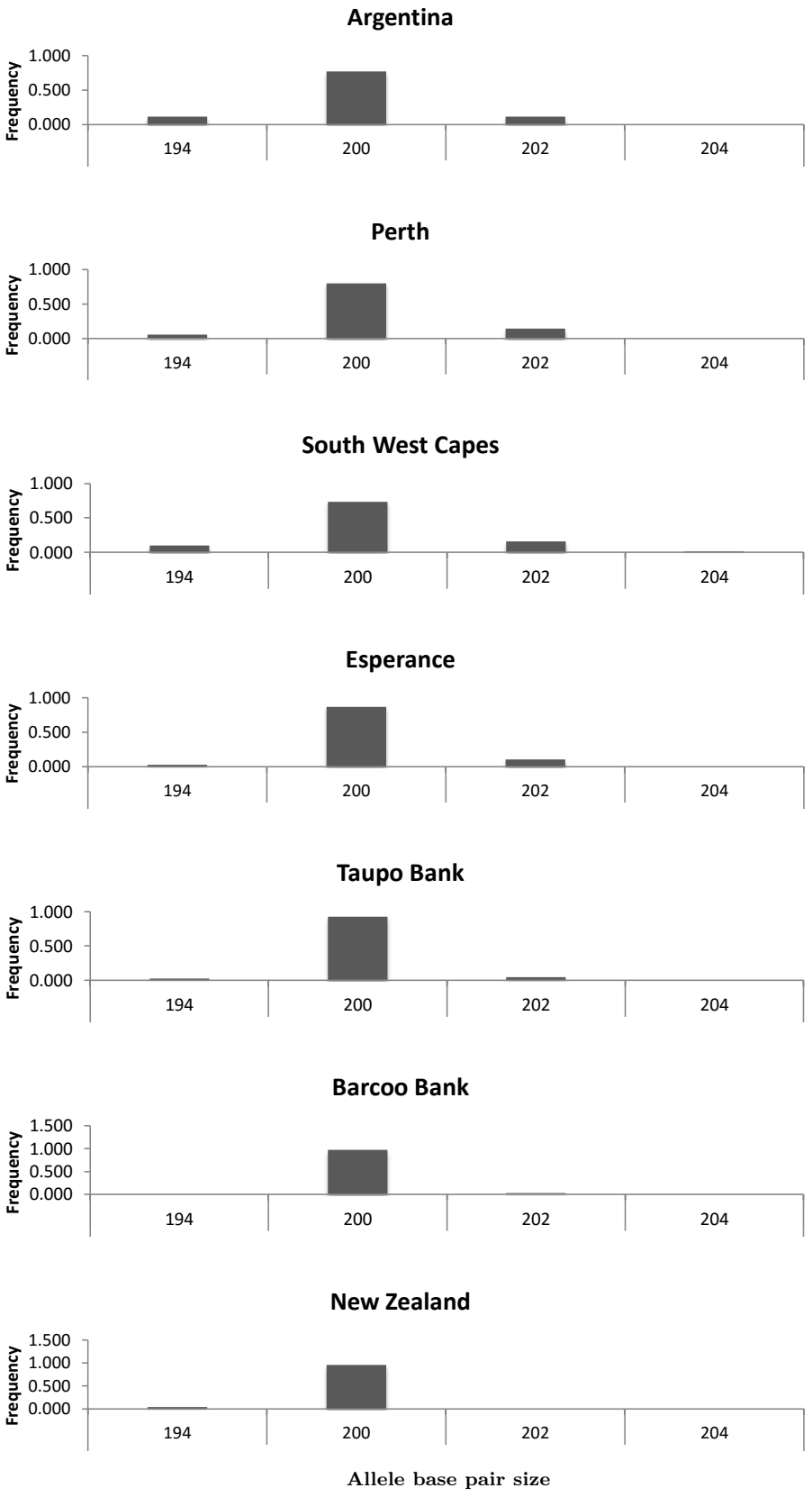



FIGURE B.6: Allelic Frequency Distribution for the microsatellite locus PAM017. Allele sizes are given in base pairs of PCR products.

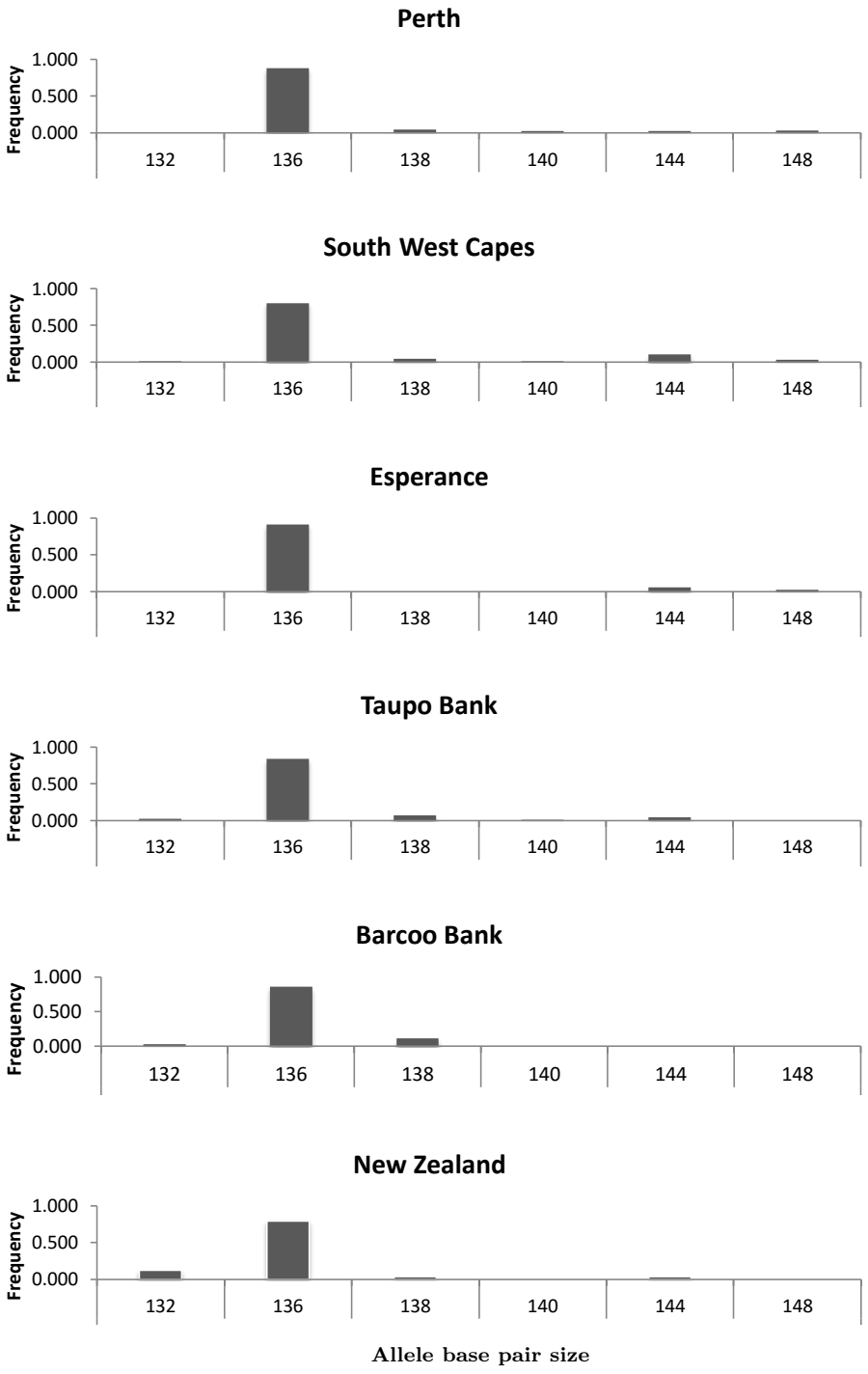



FIGURE B.7: Allelelic Frequency Distribution for the microsatellite locus PAM025. Allele sizes are given in base pairs of PCR products.

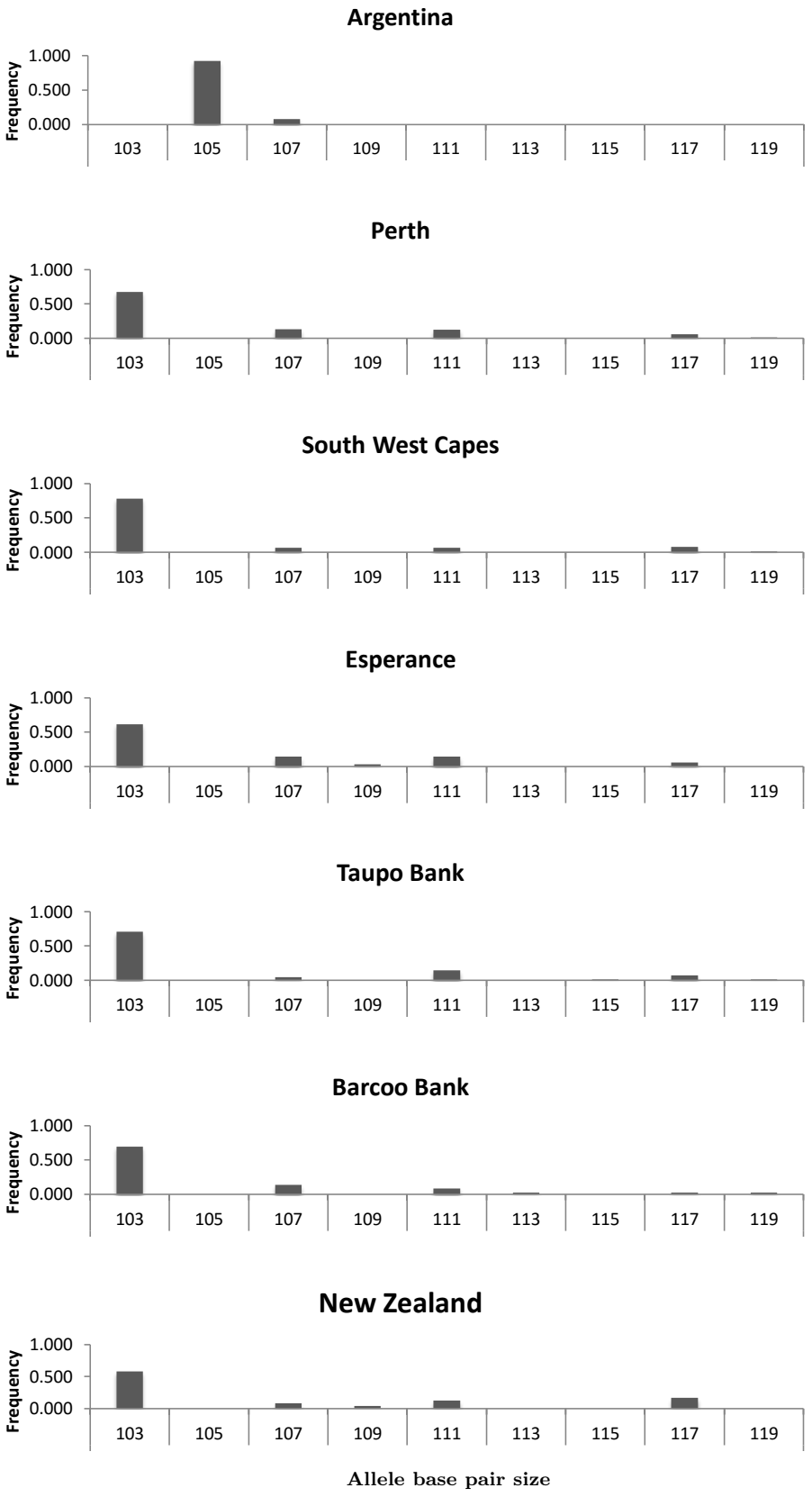



FIGURE B.8: Allelic Frequency Distribution for the microsatellite locus PAM035. Allele sizes are given in base pairs of PCR products.

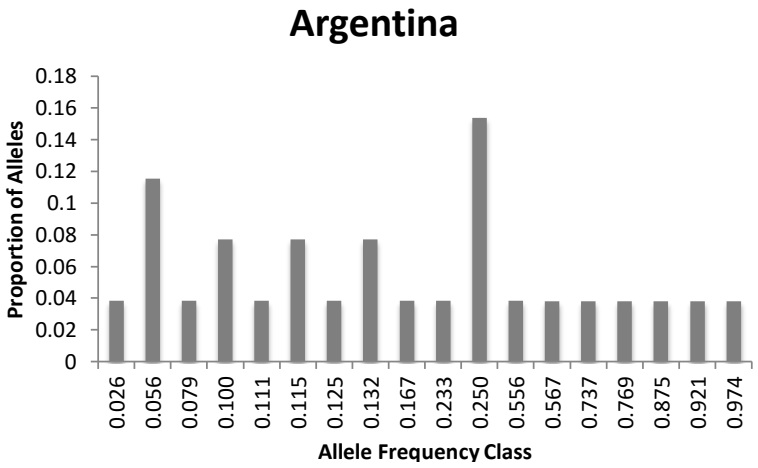



FIGURE B.9: Rare allele graph for Argentina.

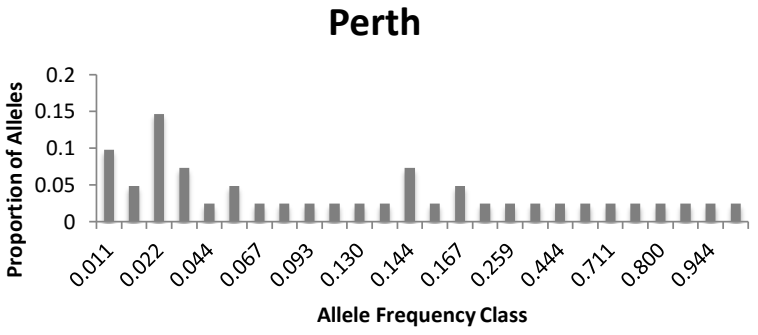



FIGURE B.10: Rare allele graph for Perth.

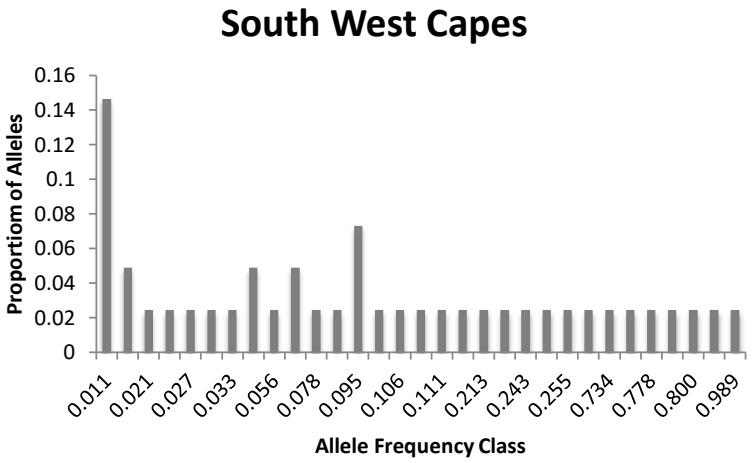



FIGURE B.11: Rare allele graph for South West Capes.

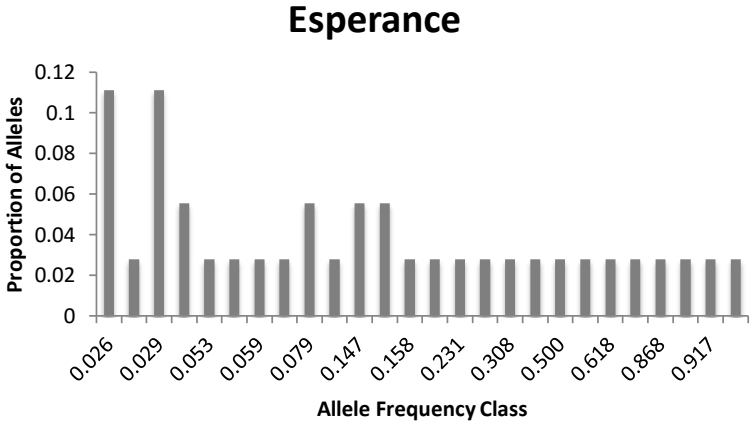



FIGURE B.12: Rare allele graph for Esperance.

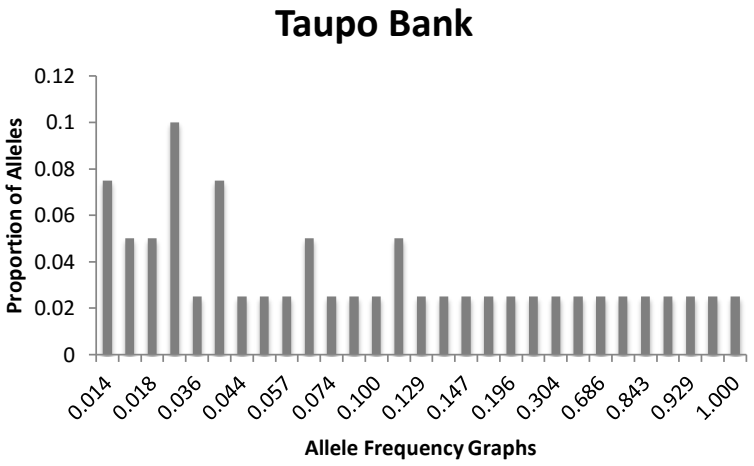



FIGURE B.13: Rare allele graph for Taupo Bank.

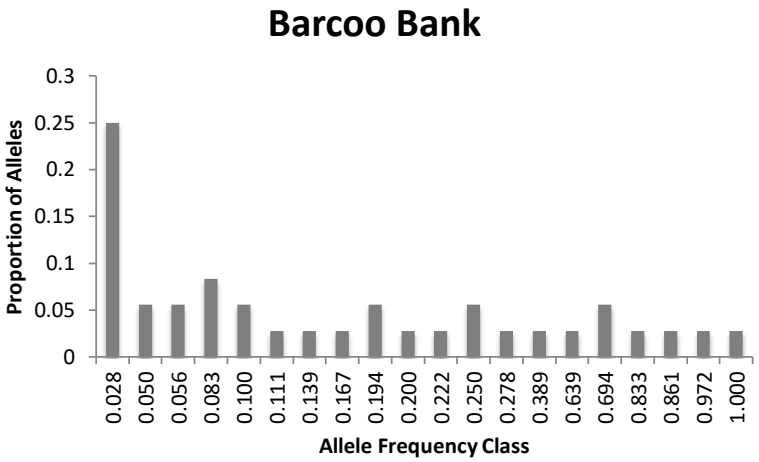



FIGURE B.14: Rare allele graph for Barcoo Bank.

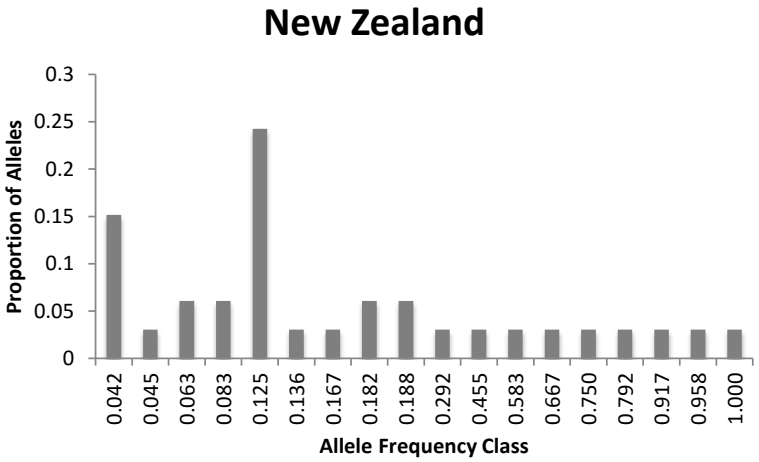



FIGURE B.15: Rare allele graph for New Zealand.

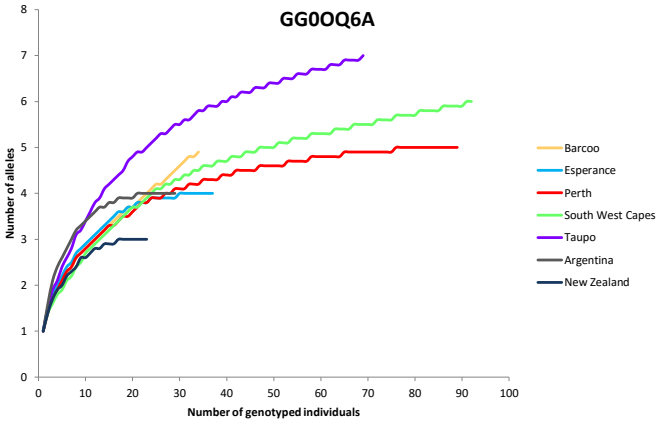



FIGURE B.16: Rarefaction curve for the locus GG0OQA.

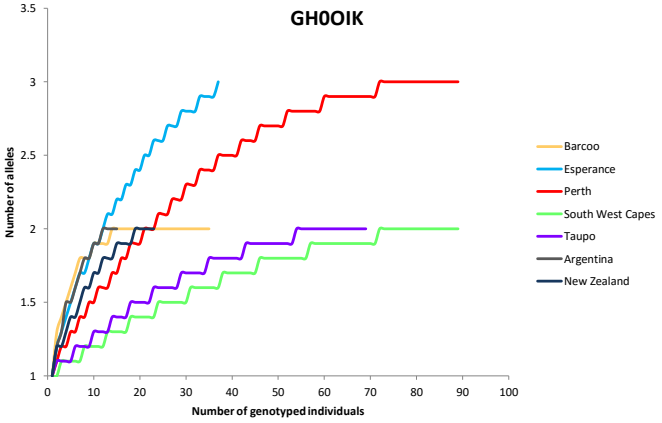



FIGURE B.17: Rarefaction curve for the locus GH0OIK.

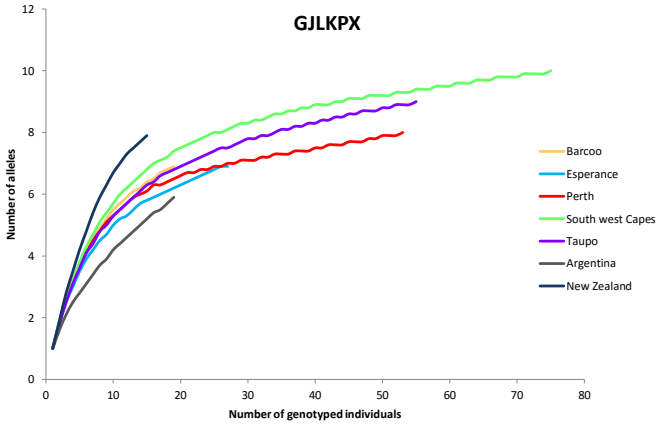



FIGURE B.18: Rarefaction curve for the locus GJLKPX.

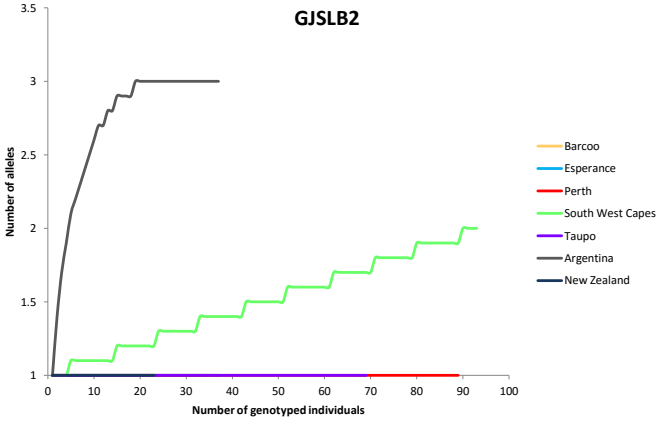



FIGURE B.19: Rarefaction curve for the locus GJSLB2.

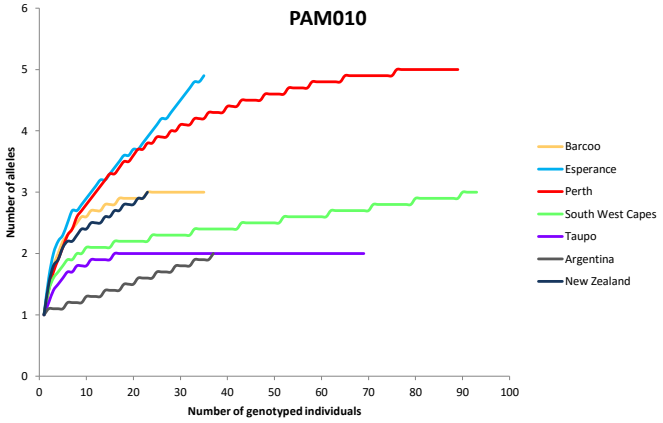



FIGURE B.20: Rarefaction curve for the locus PAM010.

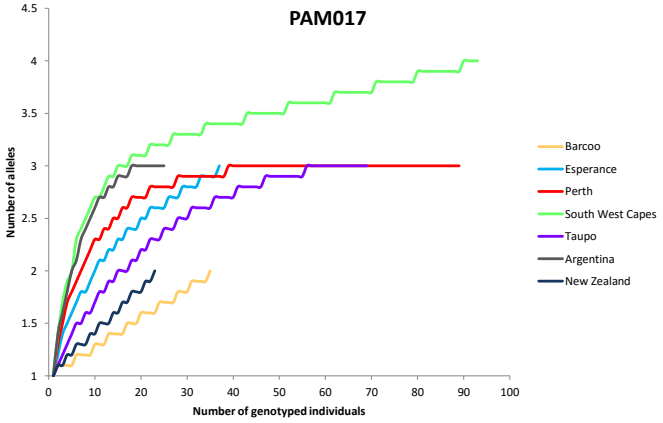



FIGURE B.21: Rarefaction curve for the locus PAM017.

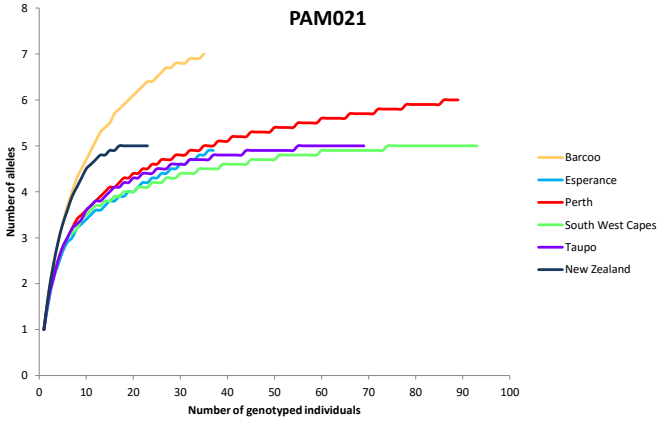



FIGURE B.22: Rarefaction curve for the locus PAM021.

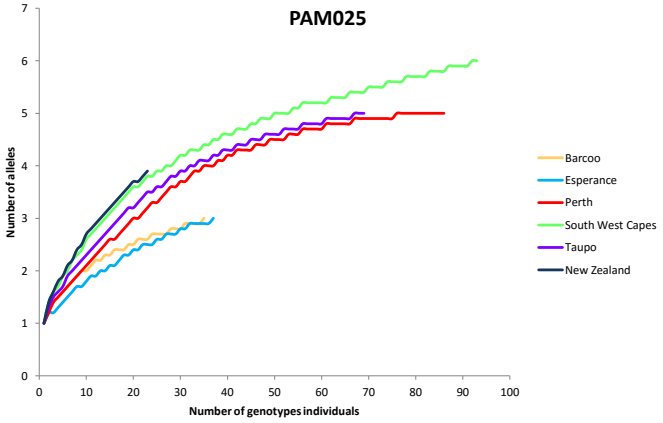



FIGURE B.23: Rarefaction curve for the locus PAM025.

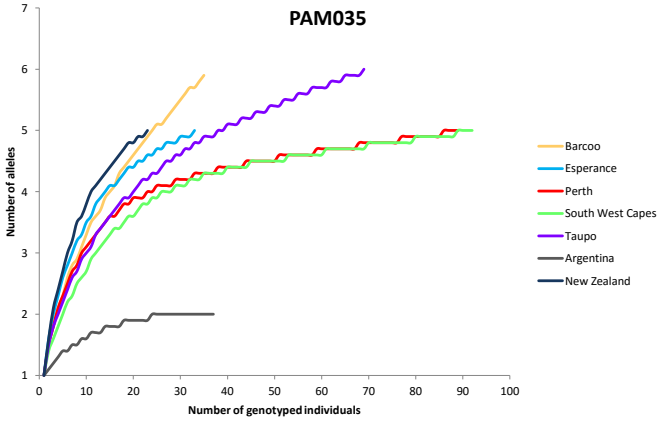



FIGURE B.24: Rarefaction curve for the locus PAM035.

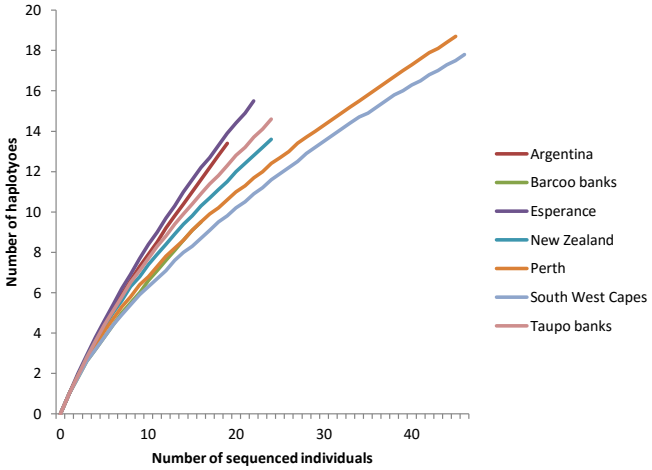



FIGURE B.25: Rarefaction curve for the mtDNA analysis in *P. americanus*.

Table B.1: G'ST and Jost's D results excl. 3 loci. G'st levels below with jost's above

| | AB | AE | AP | AS | AT | AAR | N |
|---|---|---|---|---|---|---|---|
| AB | | –0.001 | –0.005 | 0.006 | 0.006 | 0.59 | –0.013 |
| AE | –0.001 | | 0.01 | 0.03 | 0.04 | 0.631 | –0.001 |
| AP | –0.007 | 0.012 | | –0.002 | 0.003 | 0.594 | –0.009 |
| AS | 0.009 | 0.038 | –0.003 | | 0.008 | 0.605 | 0.004 |
| AT | 0.01 | 0.054 | 0.004 | 0.013 | | 0.555 | –0.001 |
| AAR | 0.477 | 0.469 | 0.465 | 0.48 | 0.481 | | 0.581 |
| N | –0.018 | –0.001 | –0.012 | 0.006 | –0.001 | 0.463 | |

Table B.2: Results from the Wilcoxon bottleneck test for *P. americanus*

| | | | IAM | | | TPM | | | SMM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Population | Loci | N | H0 | He | p | H0 | He | p | H0 | He | p |
| PER | 9 | 86 | 3 | 4.49 | 0.24 | 2 | 4.67 | 0.06 | 1 | 4.76 | 0.01 |
| SWC | 9 | 92 | 3 | 4.8 | 0.19 | 3 | 5 | 0.16 | 2 | 5.01 | 0.04 |
| ESP | 9 | 37 | 3 | 4.54 | 0.23 | 1 | 4.72 | 0.01 | 0 | 4.8 | 0 |
| BAR | 9 | 34 | 4 | 4.32 | 0.55 | 4 | 4.48 | 0.5 | 3 | 4.62 | 0.21 |
| TAU | 9 | 68 | 3 | 4.35 | 0.27 | 3 | 4.41 | 0.26 | 0 | 4.51 | 0.001 |
| NZL | 9 | 23 | 3 | 4.39 | 0.26 | 3 | 4.71 | 0.19 | 2 | 4.4 | 0.09 |

IAM=infinate alleles model, TPM=two phase model, SMM=stepwise mutation model, N=sample size (N), H0=observed heterozygosity, He=expected heterozygosity, and p=p-value.

# References

[1] Aboim ,M.A., Menezes, G.M., Schlitt, T., Rogers, A.D. (2005) Genetic structure and history of populations of the deep-sea fish *Heliolenus dactylopterus* (Delaroche, 1809) inferred from mtDNA sequence analysis. *Molecular Ecology* 14: 1343-1354.

[2] Allendorf, F. P. England, G. Luikart, Ritchie, P. (2008) Genetic effects of harvest on wild animal populations. *Trends in Ecology and Evolution*, p. 327-337.

[3] Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A., Luikart, G. (2008) LOSITAN: A workbench to detect molecular adaptation based on F(st)-outlier method. *Bmc Bioinformatics* 9:5.

[4] Ball, A.O., Sedberry, G.R., Zatcoff, M.S., Chapman, R.W., Carlin, J.L., (2000) Population structure of the wreckfish *Polyprion americanus* determined with microsatellite genetic markers. *Marine biology* 137: 1077-1090.

[5] Barreiros, J.P., Machado, L., Hostim-Silva, M., Sazima, I., and Heemstra, P.C. (2004) First record of *Polyprion oxygeneios* (Perciformes: Polyprionidae) for the south-west Atlantic and a northernmost range extension. *Journal of Fish Biology* 64(5): 1439-1441.

[6] Beaumont, M.A., Nichols, R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of Biological Sciences* 263: 1619-1626.

[7] Beckley, L.E., Muhling, B.A., Gaughan, D.J. (2009) Larval fishes off Western Australia: influence of the Leeuwin Current. *Journal of the Royal Society of Western Australia* 92: 101-109.

[8] Beentjes, M.P. and Francis, M.P. (1999) Movement of hapuku (*Polyprion oxygeneios*) determined from tagging studies. *New Zealand Journal of Marine and Freshwater Research* 33: 1-12.

[9] Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F. (2000) GENETIX 4.02, logiciel sous Windows TM pour la genetique des populations., Laboratoire Génome, Populations, Interactions, CNRS UPR 9060, Université de Montpellier II. Montpellier, France.

[10] Broquet, T., Viard, F., Yearsley, J.M., (2012) Genetic drift and collective dispersal can result in chaotic genetic patchiness. *Evolution* 67:1660-1675.

[11] Brown, W. M., George Jr., M., Wilson, A. C. (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America* 76: 1967 – 1971.

[12] Bruce, B.D., Condie, S.A., Sutton, C.A. (2001) Larval distribution of blue grenadier (*Macruronus novaezelandiae* Hector) in south-eastern Australia: further evidence for a second spawning area. *Marine and Freshwater Research* 52:603-610.

[13] Carvalho, G.R., Hauser, L. (1995) Molecular genetics and the stock concept in fisheries. *Molecular Genetics in Fisheries* 55-79.

[14] Chapuis, P., Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24(3): 621-631.

[15] Chauhan, T., Rajiv, K. (2010) Molecular markers and their applications in fisheries and aquaculture. *Advances in Bioscience and Biotechnology* 1:281-291.

[16] Ciannelli, L., Fisher, J.A.D., Skern-Mauritzen, M., Hunsicker, M.E., Hidalgo, M., Frank, K.T., Bailey, K.M. (2013) Theory, consequences and evidence of eroding population spatial structure in harvester marine fishes: a review. *Marine Ecology Progress Series* 480: 227-243.

[17] Condie, S.A., Mansbridge, J.V., Cahill, M.L. (2011) Contrasting local retention and cross-shore transports of the East Australian Current and the Leeuwin Current and their relative influences on the life histories of small pelagic fishes. *Deep-Sea Research* (II)58:606-615.

[18] Cowen, R.K., and Sponaugle, (2009) Larval dispersal and marine population connectivity. *Annual Review of Marine Science: V.1: Palo Alto, Annual Reviews*, p.443-466.

[19] Dammannagoda, S.T., Hurwood, D.A., Mather, P.B. (2008) Evidence for fine geographical scale heterogeneity in gene frequencies in yellowfin Tuna (*Thunnus albacares*) from the north Indian Ocean around Sri Lanka. *Fisheries Research* 90: 147-157.

[20] DeWoody, J.A., and J.C. Avise, 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology* 56: 461-473.

[21] DiBattista, J.D., Rocha, L.A., Craig, M.T., Feldheim, K.A., Bowen, B.W. (2012) Phylogeography of two closely related Indo-Pacific butterflyfishes reveals divergent evolutionary histories and discordant results from mtDNA and microsatellites. *Journal of Heredity* 103(5): 617-629.

[22] Donaldson, K.A., and Raymond, R., Wilson (1999) Amphi-Panamic geminates of Snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Molecular Phylogenetics and Evolution* 13(1): 208-213.

[23] Drummond, A.J., M.A. Suchard, D. Xie, and A. Rambout, 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7: *Molecular Biology and Evolution* 29: 1969-1973.

[24] Earl, D.A., and B.M. VonHoldt, 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4: 359-361.

[25] Elliot, N.G., Lowry, P.S., Grewe, P.M., Innes, B.H., Yearsley, G.K., Ward, R.D (1998) Genetic evidence for depth- and spatially separated stocks of the deep-water spikey oreo in Australasian waters. The*Fisheries Society of the British Isles* 52: 796-816

[26] Estoup, A., Jarne, P., Cornuet, J.M. (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology* 11: 1591-1604.

[27] Excoffier, L., Laval, G., Schneider, S. (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1: 47-50.

[28] Fairclough, D.V., Molony, B.W., Crisafulli, B.M., Keay, I.S., Hesp, S.A. and Marriot, R.J. (2014) Status of demersal finfish stocks on the west coast of Australia. Fisheries Research Report No. 253.

[29] Fletcher, W.J., Santoro, K.(eds). (2015) Status reports of the fisheries and aquatic resources of Western Australia 2014/15: The State of the Fisheries. Department of Fisheries, Western Australia.

[30] Fowler, A.J., Black, K.P., Jenkins, G.P. (2000) Determination of spawning areas and larval advection pathways for King George whiting in south-eastern Australia using otolith microstructure and hydrodynamic modelling. II. South Australia. *Marine Ecological Progress Series* 199: 243-254.

[31] Francis, M.P., Mulligan, K.P., Davies, N.M. and Beentjes, M.P. (1999) Age and growth estimates for New Zealand hapuku, *Polyprion oxygeneios. Fishery Bulletin* 97: 227-242.

[32] Friess C, Sedberry GR (2011) Genetic evidence for a single stock of the deep-sea teleost *Beryx decadactylus* in the North Atlantic Ocean as inferred from mtDNA control region analysis. *Journal of Fisheries Biology* 78: 466–478

[33] Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915-925.

[34] Gaither, M. R., Jones, S. A., Kelley, C., Newman, S. J., Sorenson, L., Bowen, B. W. (2011) High connectivity in the deepwater snapper *Pristipomoides filamentosus* (Lutjanidae) across the Indo-Pacific with isolation of the Hawaiian archipelago. *PLoS ONE* 6: e28913

[35] Gordon, A.L. (1971) Oceanography of Antarctic waters. In: Reid JL (ed) Antarctic oceanology. I. American Geophysical Union, Washington DC: 169-203.

[36] Goudet, J. (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* 86: 485-486.

[37] Harpending, H.C. (1994) Signature of ancient population growth on a low resolution mitochondrial DNA mismatch distribution. *Human Biology* 66: 591-600.

[38] Hauser, L., Adcock, G.J., Smith, P., Bernal Ramirez, J.H., Carvalho, G.R. (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *PNAS* 99(18): 11742-11747.

[39] Hauser, L, and G.R. Carvalho, (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries* 9: 333-362.

[40] Hauser, L., Seeb, J.E. (2008) Advances in molecular technology and their impact on fisheries genetics. *Fish and Fisheries* 9: 473-486.

[41] Hedgecock, D., Barber, P.H., Edmonds, S. (2007) Genetic approaches to measuring connectivity. *Oceanography* 20:70–79.

[42] Hedrick, P.W. (2005) A standardized genetic differentiation measure. *Evolution* 59: 1633-1638.

[43] Hellberg, M.E. (2009) Gene flow and isolation among populations of marine animals. Annual Review of Ecology, Evolution and Systematics. *Palo Alto, Annual Reviews* 40: 291-310.

[44] Henriques, R., Potts, C.V., Santos, C.V., Sauer, W.H.H., Shaw, P.W. (2014) Population connectivity and phylogeography of a coastal fish, *Atractoscion aequidens* (Sciaenidae), across the Benguela Current Region: Evidence of an Ancient Vicariant Event. *Plos One* 9.

[45] Ho, S.Y.W., Shapiro, B. (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources* 11: 423-434.

[46] Hoffman, J.L., Dasmahapatra, K.K., Amos, W., Phillips, C.D., Gelatt, T.S., Bickham, J.W. (2009) Contrasting patterns of genetic diversity at three different genetic markers in a marine mammal population. *Molecular Ecology* 18: 2961-2978.

[47] Holland, S.M. (2003) Analytic Rarefaction 1.3, http://strata.uga.edu/ software/index.html.

[48] Holsinger, K.E., Weir, B.S (2009) Genetics in geographically structured populations: defining, estimating and interpreting Fst. *Nature Reviews* 10: 639-649.

[49] Imbrie, J., Boyle, E.A., Clemens, S.C., Duffy, A., Howard, W.R., Kukla, G., Kutzbach, J., Martinson, D.G., McIntyre, A., Mix, A.C., Molfino, B (1992) On the structure and origin of major glaciation cycles 1. Lnear responses to Milankovitch forcing. *Paleooceanography* 7(6): 701-738.

[50] Jenkins, G.P., Black, K.P., Hamer, P.A. (2000) Determination of spawning areas and larval advection pathways for King George whiting in southe-astern Australia using otolith microstructure and hydrodynamic modelling. I. Victoria. *Marine Ecological Progress Series* 199: 231-242.

[51] Jost, L., (2008) Gst and its relatives do not measure differentiation. *Molecular Ecology* 17: 4015-4026.

[52] Kirk, H., and Freeland, J.R. (2011) Applications and implications of neutral versus non-neutral markers in molecular ecology. *International Journal of Molecular Sciences* 12: 3966-3988.

[53] Knutsen H, Jorde P.E, Sannaes H, Hoelzel A.R, Bergstad O.A, Stefanni S, Johansen T, Stenseth N.C (2009) Bathymetric barriers promoting genetic structure in the deepwater demersal fish tusk (*Brosme brosme*). *Molecular Ecology* 18:3151–3162.

[54] Knutsen, H., Olsen, E.M., Jorde, P.E., Espeland, S.H., Andre, C., Stenseth, N.C. (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: 768-783.

[55] Laikre, L., Palm, S., Ryman, N. (2005) Genetic population structure of fishes: implications for coastal zone management. *A Journal of the Human Environment* 34: 111-119.

[56] Lane, H.S. (2013) Characterisation of the mitochondrial genome and the population genetics of *Polyprion oxygeneios* (hapuku) from around New Zealand. Victoria University of Wellington.

[57] Lane, H.S., Symonds, J.E, Ritchie, P.A (2016) The phylogeography and population genetics of *Polyprion oxygeneios* based on mitochondrial DNA sequences and microsatellite DNA markers. *Fisheries Research* 174: 19-29.

[58] Librado, P., Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

[59] Lowe, W.H., and F.W. Allendorf, 2010. What can genetics tell us about population connectivity?. *Molecular Ecology* 19: 3038-3051.

[60] Machias, A., Somarakis, S., Karakassis, I., Neofytou, C., Maravelias, X., Pantazis, P. (2003) Fish landing changes since the onset of aquaculture: two case studies. *Abstracts 7th Hellenic Symposium Oceanography and Fisheries, Iraklion* p. 203 ISSN 1107-6534.

[61] McCusker, M.R., Bentzen, P. (2010) Positive relationships between genetic diversity and abundance in fishes. *Molecular Ecology* 19: 4852–4862.

[62] Meirmans, P.G., Van Tienderen, P.H. (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792-794.

[63] Ministry for Primary Industries (MPI) (2012). Fisheries Assessment Plenary, November 2012: stock assessments and yield estimates. Compiled by the Fisheries Science Group, Ministry for Primary Industries, Wellington, New Zealand.

[64] Ministry for Primary Industries (MPI) (2015). Fisheries Assessment Plenary, May 2015: Stock Assessments and Stock Status. Compiled by the fisheries Science Group, Ministry for Primary Industries, Wellington, New Zealand.

[65] Nielsen, E.E., Hemmer-Hanson, J.H., Poulsen, N.A., Loeschcke, V. (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology* 9: 276.

[66] Nielsen, J.L., Graziano, S.L., Seitz, A.C. (2010) Fine-scale population genetic structure in Alaskan Pacific halibut (*Hippoglossus stenolepis*). *Conservation Genetics* 11: 999–1012.

[67] Olsen, E.B., Knutsen, H., Gjosaeter, J., Jorde, P.E., Knutsen, J.A., Stenseth, N.C. (2008) Small-scale biocomplexity in coastal Atlantic cod supporting a Darwinian perspective on fisheries management. *Evolutionary Applications* 1(3): 524-533.

[68] Ovenden , J., 2013. Crinkles in connectivity: combining genetics and other types of biological data to estimate movement and interbreeding between populations. *Marine and Freshwater Research* 64: 201-207.

[69] Paul, L.J. (2002) Can existing data describe the stock structure of the two New Zealand groper species, hapuku (*Polyprion oxygeneios*) and bass (*P.americanus*)? *New Zealand Fisheries Assessment Report* 2002/14.

[70] Peakall, R., Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research: *Molecular Ecology Notes* 6: 288-295.

[71] Peakall, R. and Smouse, P.E. (2012) Genalex 6.5: genetic analysis in Excel. Population genetic software for teaching and research- an update. *Bioinformatics* 28: 2537-2539

[72] Peres, M.B. and Klippel, S. (2003) Reproductive biology of the south-western Atlantic wreckfish *Polyprion americanus* (Teleosti: Polyprionidae). *Environmental Biology of Fishes* 68: 163-173.

[73] Peres, M.B., Haimovici, M. (2004) Age and growth of south-western Atlantic wreckfish *Polyprion americanus*. *Fisheries Research* 66(2-3): 157-169.

[74] Piry, S., Luikart, G., Cornuet, J.M. (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* 90: 502-503.

[75] Pritchard, J.K., Stephens, M., Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

[76] Rambout A., Suchard, M., Xie, D., Drummond, A. (2014) Tracer v1.6, http://beast.bio.ed.ac.uk/Tracer

[77] Raymond, M., and F. Rousset (1995) GENEPOP (VERSION 1.2)- Population genetics software for exact test and ecumenicism. *Journal of Heredity* 86: 48-249.

[78] Robert, C.D. (1986) Systematics of the percomorph fish genus *Polyprion* Oken, 1917. PhD dissertation. Victoria University of Wellington, Wellington, New Zealand.

[79] Rogers, A.R., Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9: 552-569.

[80] Ross, P.M., I.D. Hogg, C.A. Pilditch, and C.J. Lundquist, (2009) Phylogeography of New Zealand's coastal benthos. *New Zealand Journal of Marine and Freshwater Research* 43: 1009-1027.

[81] Ruzzante, D.E., Taggart, T.E., Cook, D. (1998) A nuclear DNA basis for shelf- and bank-scale population structure in north-west Atlantic cod (*Gadus morhua*): Labrador to Georges Bank. *Molecular Ecology* 7: 1663-1680.

[82] Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) Molecular cloning: a lab manual v. 3: New York, Cold Spring Harbour Laboratory Press.

[83] Sedberry, G.R., Carlin, J.L., Chapman, R.W., Eleby, B. (1996) Population structure of the pan-ceanic wreckfish, *Polyprion americanus* (Teleosti: Polyprionidae), as indicated by mtDNA variation. *Journal of fish Biology* 49(Suppl A): 318-329.

[84] Sedberry, G.R., Andrade, C.A.P., Carlin, J.L., Chapman, R.W., Luckhurst, B.E., Manooch, C.S., Meneses, G., Thomsen, B., Ulrich, G.F. (1999) Wreckfish *Polyprion americanus* in the North Atlantic: Fisheries, biology, and management of a widely distributed and long-lived fish. *American Fisheries Society Symposium* 23: 27-50.

[85] Selkoe, K.A., Toonen, R.J (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9(5): 615-629.

[86] Schloetterer, C. (2004) The evolution of molecular markers- just a matter of fashion? *Nature Reviews Genetics* 5: 63-69.

[87] Scoles, D.R., Collette, B.B., Graves, J.E. (1998) Global phylogeography of mackerels of the genus *Scomber*. *Fish Bulletin* 98: 823-842.

[88] Shaw, P.W., Turan, C., Wright, J.M., O'Connell, M., Carvalho, G.R., (1999) Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. *Heredity* 83: 490-499.

[89] Smallwood, C.B., Hesp, S.A., and Beckley, L.E. (2013) Biology, stocks status and management summaries for selected fish species in south-western Australia. Fisheries Research Report No. 242.

[90] Smith, P.J., Johnson, A.D. (1985) Glucosephosphate isomerase and $\alpha$-glycerophosphate dehydrogenase electromorph frequencies in groper (*Polyprion oxygeneios*)from central New Zealand. *New Zealand Journal of Marine and Freshwater Research* 19: 173-177.

[91] Smith, P. J. (2008) Population biology and genetics of paua, hapuku and kingfish: sourcing fish for broodstock development. *NIWA Technical Report 132*. 44 p.

[92] Stepien, C.A., Dillon, A.K., Patterson, A.K. (2000) Population genetics, phylogeography, and systematic of the thornyhead rockfishes (*Sebastolobus*) along the deep continental slopes of the North Pacific Ocean. *Canadian Journal of Fisheries and Aquatic sciences* 57: 1701-1717.

[93] Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA Polymorphism. *Genetics* 123: 585-595.

[94] Takeshima, H., Hatanaka, A., Yamada, S., Yamazaki, Y., Kimura, I., Nishida, M. (2011) Low genetic differentiation between two geographically separated populations of demersal gadiform fishes in the Southern Hemisphere. *Genes and Genetic Systems* 86: 339-349.

[95] Tatarenkov, A., Joensson, R.B., Kautsky, L., Johannesson, K. (2007) Genetic structure in population of *Eucus vesiculopsus* (Phaeophuceae) over spatial scales from 10m to 800km. *Journal of Phycology* 43: 675-685.

[96] Teacher, A.G.F., Andre, C., Jonsson, P.R., Merila, J. (2013) Oceanographic connectivity and environmental correlates of genetic structuring in Atlantic herring in the Baltic Sea. *Evolutionary Applications* 6: 549-567.

[97] Van Oosterhout, C., Hutchinson, W.F., Willis, D.P.M., Shiply, P. (2004) MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538.

[98] Varela, A.I., Ritchie, P.A., Smith, P.J. (2013) Global genetic population structure in the commercially exploited deep-sea teleost Orange Roughy (*Haplostethus atlanticus*) based on microsallite DNA analyses. *Fisheries Research* 140: 83-90.

[99] Vigliola, L., Doherty, P.J., Meekan, M.G., Drown, D.M., Jones, M.E., Barber, P.H. (2007) Genetic identity determines risk of post-settlement mortality of a marine fish. *Ecology* 88: 1263-1277.

[100] Von der Heyden, S., Lipinski, M.R., Matthee, C.A. (2007) Mitochondial DNA analyses of the Cape hakes reveal an expanding, panmictic population for *Merlucciuc paradoxus*. *Molecular Phylogenetic Evolution* 42: 515-527.

[101] Waite, A.M., Thompson, P.A., Feng, M., Beckley, L.E., Dominigues, C.M., Gaughan, D., Hanson, C.E., Holl, C.M., Koslow, T., Meuleners, M., Montoa, J.P., Moore, T. (2007) The Leeuwin Current and its eddies: an introductory overview. *Deep Sea Research Part II: Topical Studies in Oceanography* 54(8-10): 789-796.

[102] Wakefield, C.B., Newman, S.J., Molony, B.W. (2010) Age-based demography and reproduction of hapuku, *Polyprion oxygeneios*, from the south coast of Western Australia: implications for management. *ICES Journal of Marine Science* 67: 1164-1174.

[103] Waples, R.S. (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *The Journal of Heredity* 89: 438-450.

[104] Waples, R.S., Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15: 1419-1439.

[105] Ward, R.D. (2000) Genetics in fisheries management. *Hydrobiologica* 420: 191-201.

[106] Weir, B.S., Cockerham, C.C. (1984) Estimating f-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

[107] Wennerstroem, L., Laikre, L., Ryman, N., Utter, F.M., Ghani, N.I.A., Andre, C., DeFaveri, J., Johansson, D., Kautsky, L., Merila, J., Mikhailova, N. (2013) Genetic biodivesity in the Baltic Sea: species-specific patterns challenge management. *Biodiversity Conservation* 22: 3045-3065.

[108] Wise, B.S., St John, J., Lenanton, R.C.J. (2007) Spatial scales of exploitation among populations of demersal scalefish: implications for management. 1. Stock status of the key indicator species for the demersal scalefish fishery in the west coast bioregion. Final report to fisheries research and development corporation

on project No. 2003/052. Department of Fisheries, Western Australia. Fisheries Research Report 163. 130 pp.