

Population genetics and spawning time of Lake Taupo rainbow trout

Elizabeth Rose Heeg

A thesis submitted to Victoria University of Wellington in fulfilment of the requirement
for the degree of Doctor of Philosophy in Ecology and Biodiversity

Victoria University of Wellington

Te Whare Wānanga o te Ūpoko o te Ika a Māui

2012

Abstract

The rainbow trout (*Oncorhynchus mykiss*) of Lake Taupo, New Zealand provide an exceptional opportunity to explore the contemporary adaptation of an introduced aquatic species. Recently it has become evident that their spawning migration time has shifted to later in the season. I investigated the genetic basis of these changes in spawning time by (1) using genetic markers to determine the origins of Taupo trout in California, (2) determining the pattern and extent of spatial population genetic variation throughout the Lake Taupo catchment and in comparison to nearby Lake Tarawera in the Rotorua district, (3) analysing genetic variation at the *OtsClock1b* spawning time gene in temporal replicates from several sites from Taupo, and (4) comparing contemporary genetic variation at this gene and microsatellite markers to genetic variation from three Taupo tributaries in 1980s. I compared the ability of single nucleotide polymorphism (SNP) and microsatellite markers to determine the origins of Lake Taupo rainbow trout, translocated from California around 120 years ago. Data were collected from 15 microsatellite and 93 SNP markers, using samples from the Lake Taupo population and ten populations throughout California, which included all historically indicated populations of origin. Results revealed that the Lake Taupo population has significantly diverged from Californian populations at both microsatellite and SNP loci. These analyses also showed that the Lake Taupo population was probably derived from several sources in California (the most likely origins being the McCloud River and Lake Almanor), and an indeterminate California coastal population. This conclusion was supported with simulations of founder events, which suggested that the genetic patterns of a single source of introduction would still be detectable 100 years post-founding, but with multiple introductions exact source populations become more difficult to detect. Approximately 50 individuals from 10 locations throughout the catchment were then analysed using 15 microsatellite loci to determine if there was any spatial population genetic differentiation. There was no significant difference in genetic distance between locations within Lake Taupo, although there was a significant difference between these populations and Rotorua and Waipakihi, which are isolated by geographic barriers. Lake Taupo rainbow trout do appear to diverge at markers potentially under selection, though, because genotyping of the poly-Q region of the timing locus *OtsClock1b* shows significant differentiation between individuals sampled at different times in the Waipa River. Two other sites, however, did not show the same pattern of significant seasonal variation in *OtsClock1b*

allele frequencies. This suggests that genotypes at this locus could be influencing spawning migration time, but that this variation could also be site specific, and therefore have a strong environmental component. Scale samples from the 1980s show no significant divergence at 5 microsatellites and *OtsClock1b*, indicating that allele frequencies have not changed significantly over the last 20 years at neutral markers or markers under selection. I therefore conclude that while Taupo rainbow trout have diverged from their origins in California, they have only slightly diverged within their new environment, and do not show a consistent pattern of genetic change over time. This information will contribute not only to the management of the Taupo fishery but also to the current understanding of the population genetic structuring of introduced salmonids.

*This work is dedicated to Charles and Violet Rose, whose
commitment to education left an enduring legacy for their family.*

Acknowledgements

Thanks to Dr. Peter Ritchie for his support and guidance through the last three years. I also appreciate the support from Dr. Phil Lester as my secondary supervisor and Head of School for the latter part of my PhD. Many thanks to Dr. Fred Allendorf for inviting me to participate in this project in the first place, and for his contributions to the work contained herein.

Thanks to Cameron Jack, Ange Fleming, and Chris Thorn for technical support at Victoria University. Big thanks also go to Sandra Taylor, Mary Murray, and Paul Marsden for administrative support and a sneaky glass of wine every now and again.

This work would not have been possible without the dedicated staff of the Taupo-nui-a-Tia Department of Conservation office, particularly Dr. Michel Dedual and Glenn Maclean. Their financial and technical support gave me the opportunity to do interesting work in a beautiful place, and I appreciate how rare chances like that are.

Thanks also to the Massey Genome Service for genotyping of samples.

This project was supported a grant from the Department of Conservation. I was personally supported by a Victoria University of Wellington PhD scholarship and a Victoria University of Wellington PhD Submission scholarship. My travel to California for work presented in Chapter 2 was supported by a New Zealand Educated Travel Award. My travel to the Coastwide Salmon meeting in Olympia, WA, USA was supported by an award from the American Fisheries Society Genetics Section.

Acknowledgements by chapter:

Chapter 1: Thanks to Danielle Hannan for editing assistance. Thanks to the late Dr. Bob McDowall who sent me a copy of his excellent book on acclimatization societies.

Chapter 2: This chapter was co-authored as a paper for publication with Dr. J. Carlos Garza, Dr. Sebastien Rioux Paquette, Dr. Michel Dedual, Alicia Abadía-Cardosa, and Dr. Peter Ritchie. Thanks to Glenn Maclean, John Gibbs, and the staff at the Department of Conservation Office in Turangi for sampling and helpful discussions; Dr. Fred Allendorf for development and revisions; Monica Gruber, Steve Corin and Dr. Kristina Ramstad for

early revisions and discussion; and Anthony Clemento, Dr. Libby Gilbert-Horvath and the staff of the Southwest Fisheries Science Center's Molecular Ecology and Genetic Analysis Team for laboratory assistance. This work was supported by a grant from the Department of Conservation and by the Southwest Fisheries Science Center.

Chapter 3: Thanks to the Rotorua office of Fish and Game New Zealand for the Rotorua samples. Dr. Kevin Williamson assisted with revisions and discussion.

Chapter 4: Thanks to the staff at the Department of Conservation Office in Turangi for sampling and helpful discussions. Thanks to Dr. Kathleen O'Malley for her assistance with navigating the complex field of circadian rhythm genetics. Thanks also to Drs. Marc Hasenbank and Dalice Sim for assistance with statistical analysis. Monica Gruber assisted with revisions.

Chapter 5: Thanks to Dr. Marvin Rosenau for the scale samples from the 1980s. Thanks to Dr. Browyn Kivell's group for hosting me in their lab for my work on scale samples. Henry Lane enthusiastically assisted with lab work, which I greatly appreciate.

Thanks to the Ritchie lab group and the "Hatchet group" for stimulating discussions and revisions of multiple presentations of this work.

Thanks to all my friends, both in New Zealand and overseas, for their support, their sense of humour, and their high tolerance for my foolishness.

Last but not least, many thanks for my family and my partner Anrik Drenth for their enduring love and support. I owe you big.

TABLE OF CONTENTS

| | |
|---|------------------|
| <i>Abstract.....</i> | <i>iii</i> |
| <i>Acknowledgements.....</i> | <i>vii</i> |
| <i>List of figures.....</i> | <i>xi</i> |
| <i>List of tables.....</i> | <i>xiii</i> |
| Introduction..... | 1 |
| 1.1 Introduced populations | 1 |
| Freshwater fish introductions | 2 |
| 1.2 Population structure and management | 3 |
| Population management..... | 4 |
| 1.3 Rainbow trout..... | 5 |
| Background of the species | 5 |
| Introduced rainbow trout | 6 |
| The genetic changes associated with introductions..... | 7 |
| 1.4 The Lake Taupo system | 7 |
| Management of trout in New Zealand and Lake Taupo | 7 |
| Observed change in spawning time and body size..... | 8 |
| 1.5 Circadian rhythms and the genetic basis of spawn timing..... | 9 |
| Circadian rhythms..... | 9 |
| The genetic basis of spawn timing | 10 |
| 1.6 Aims of the thesis research | 11 |
| 1) Tracing the origins..... | 11 |
| 2) Spatial population genetic structure | 11 |
| 3) Spawn timing genetics..... | 12 |
| 4) Temporal genetic variation | 12 |
| <i>The origins of New Zealand rainbow trout (Oncorhynchus mykiss)</i> | <i>13</i> |
| 2.1 Introduction | 13 |
| 2.2 Materials and Methods..... | 17 |
| Sampling..... | 17 |
| DNA Extraction and Genetic Analyses..... | 17 |
| Data Analysis | 18 |
| Simulations..... | 20 |
| 2.3 Results | 21 |
| Genetic variation within populations..... | 21 |
| Genetic variation between populations..... | 21 |
| Simulations..... | 27 |
| 2.4 Discussion | 29 |
| What pattern of introduction do the data support? | 29 |
| What genetic patterns would be expected for different introduction scenarios?..... | 31 |
| Do SNP markers and microsatellites provide comparable results? | 32 |
| Conclusions | 33 |
| <i>Population genetic structure of the rainbow trout in Lake Taupo, New Zealand.....</i> | <i>35</i> |
| 3.1 Introduction | 35 |
| 3.2 Materials and Methods..... | 37 |
| Sample collection and genetic methods | 37 |
| Genetic variation | 39 |
| Spatial genetic structure..... | 39 |
| 3.3 Results | 42 |
| Genetic variation | 42 |
| Spatial genetic structure..... | 42 |

| | |
|---|------------|
| 3.4 Discussion..... | 46 |
| Genetic variation..... | 46 |
| Spatial genetic structure | 48 |
| <i>Lack of spatial variation but significant seasonal variation in OtsClock1b in Lake Taupo rainbow trout (Oncorhynchus mykiss).....</i> | 51 |
| 4.1 Introduction..... | 51 |
| 4.2 Materials and Methods..... | 54 |
| Spatial Sampling..... | 54 |
| Fishing type sampling | 55 |
| Seasonal sampling..... | 55 |
| Genomic DNA extraction and genotyping | 55 |
| Statistical analysis: Spatial variation | 57 |
| Statistical analysis: Temporal variation..... | 58 |
| Statistical analysis: Fishing type..... | 59 |
| 4.3 Results | 59 |
| OtsClock1b allele frequency distribution geographically | 59 |
| OtsClock1b allele frequency distribution seasonally | 60 |
| Effects of fishing type | 62 |
| 4.4 Discussion..... | 63 |
| OtsClock1b allele frequency distribution geographically | 63 |
| OtsClock1b allele frequency distribution seasonally | 64 |
| Effect of angling: exploitative selection? | 65 |
| <i>Minimal temporal variation in Lake Taupo rainbow trout and the importance of looking at more than one snapshot in time</i> | 71 |
| 5.2 Materials and Methods..... | 73 |
| Genomic DNA extraction and genotyping | 73 |
| Statistical analysis..... | 75 |
| 5.3 Results | 76 |
| Genetic diversity | 76 |
| Genetic differentiation..... | 79 |
| 5.4 Discussion..... | 80 |
| Genetic diversity | 80 |
| Genetic differentiation..... | 81 |
| Conclusions and future work..... | 82 |
| <i>General Discussion.....</i> | 85 |
| 6.1 Introduction..... | 85 |
| 6.2 Major findings and their implications | 86 |
| Genetic structure | 86 |
| Spawn timing genetics..... | 89 |
| Future work..... | 90 |
| 6.3 Management recommendations..... | 90 |
| Spatial genetic management..... | 90 |
| Seasonal variation: restoring the early run of Taupo rainbow trout? | 91 |
| 6.4 Contribution to the field | 92 |
| <i>Appendix.....</i> | 95 |
| <i>References.....</i> | 111 |

List of figures

| | |
|--|----|
| 2.1. Map of California showing sampling sites | 16 |
| 2.2. Distribution of allele frequencies for Lake Taupo, Lake Almanor, and the McCloud River rainbow trout populations, calculated from 15 microsatellite loci. | 25 |
| 2.3. A plot of the frequencies for all alleles in the Lake Taupo and McCloud River populations at 15 microsatellite loci. If McCloud is the sole source, there is less than a 5% chance of missing an allele present at 14% in the McCloud River during the founding event (denoted by the arrow on the x axis) | 25 |
| 2.4. A plot of microsatellite and SNP pairwise F_{ST} values for the Lake Taupo and 10 Californian populations | 26 |
| 2.5. A principal component analysis (PCA) computed from fifteen microsatellite loci and 93 SNP loci for Lake Taupo and ten Californian populations | 26 |
| 2.6. Mean F_{ST} values calculated between 1000 simulated introduced populations and ten Californian rainbow trout for multiple introduction scenarios (A-D) and empirical data (E). Results are indicated for SNP markers (SNPs), microsatellite markers (mSATs) and microsatellite markers recoded as biallelic markers (mSATsBi) | 29 |
| 3.1. Map of ten sampling locations for rainbow trout around Lake Taupo, New Zealand. | 41 |
| 3.2. A factorial correspondence analysis of the genetic differentiation of 10 riverine populations from Lake Taupo and one population from Rotorua based on allele frequencies from 14 microsatellites | 44 |
| 3.3. Comparison of geographic distance and Nei's genetic distance | 46 |
| 4.1. Historical changes in rainbow trout spawning time in the Waipa stream from 2001 to 2008, with number of rainbow trout trapped on the y-axis. Data from the Department of Conservation | 54 |

| | |
|---|-----|
| 4.2. Spatial distribution of <i>OtsClock1b</i> allele frequencies at five sampling sites around Lake Taupo. Sample size is noted at the top of the bars. | 59 |
| 4.3. Allele frequencies at <i>OtsClock1b</i> in nine Californian populations and a combined Lake Taupo sample from several different tributaries. | 60 |
| 4.4. Pairwise F_{ST} values for the November 2006 and May/June 2007 genetic samples taken at Waipa trap | 61 |
| 4.5. Allele frequencies at <i>OtsClock1b</i> compared between two different collection methods taken in winter of 2009..... | 63 |
| 4.6. Seasonal differences in allele frequencies at <i>OtsClock1b</i> at three sampling sites around Lake Taupo. Sample size is noted at the top of the bars. | 67 |
| 4.7. Seasonal differences in genotypic frequency at <i>OtsClock1b</i> at three sampling sites around Lake Taupo. | 68 |
| 4.8. Fishing effort on Lake Taupo and the Tongariro river for 2005 in hours. Figure courtesy Department of Conservation Taupo-tia-a-nui Office, Turangi | 69 |
| 5.1. A principle components analysis (PCA) of temporal samples from three Lake Taupo populations | 77 |
| 5.2. Allele frequencies of three Lake Taupo populations for the poly-Q region of <i>OtsClock1b</i> over time. Samples sizes are at the top of the bars | 77 |
| 5.3. Power analysis of the effect of sample size..... | 80 |
| A. 3.1. Allele discovery curves for ten Lake Taupo and one Rotorua population, computed by jackknife simulation | 95 |
| A. 3.2. Determination of number of K cluster from population genetic data from ten Lake Taupo populations and one Rotorua population | 99 |
| A. 4.1. R code for backwards stepwise elimination generalized linear model | 104 |

Some captions are abbreviated from those used within the text.

List of tables

| | |
|--|----|
| 2.1. Sample size (n), expected heterozygosity (H_E) and allelic richness (Ar) for Lake Taupo, New Zealand and California, USA populations computed from 15 microsatellite loci and expected heterozygosity..... | 22 |
| 2.2. Pairwise F_{ST} values from comparisons of ten <i>O. mykiss</i> populations from California and Lake Taupo, calculated from 93 SNP loci and 15 microsatellite loci. | 23 |
| 2.3. Comparison of pairwise F_{ST} values from comparisons of ten <i>O. mykiss</i> populations from California and Lake Taupo. | 24 |
| 3.1. Microsatellite loci, fluorescent dye label used for visualization, PCR multiplex, annealing temperature and PCR concentration in a 5 μ l reaction for 19 markers used in this study | 38 |
| 3.2 Sample size, expected heterozygosity (H_E), allelic richness (Ar), private alleles (PA), and inbreeding coefficient (F_{IS}) in ten Lake Taupo and one Rotorua rainbow trout populations. TT represents Tauranga Taupo | 44 |
| 3.3. Pairwise F_{ST} comparisons for ten Lake Taupo rainbow trout populations calculated from data from 14 microsatellites | 45 |
| 4.1 Samples used to investigate spatial variation in the <i>OtsClock1b</i> gene in both Lake Taupo and California | 56 |
| 4.2 Microsatellite loci, fluorescent dye label used for visualization, PCR multiplex, annealing temperature and PCR concentration in a 5 μ l reaction for 12 neutral markers used in this study..... | 57 |
| 4.3. The number (N) of male and female rainbow trout genotyped from seasonal samples at three tributaries of Lake Taupo | 58 |
| 4.4 Pairwise F_{ST} values for seasonal samples from Lake Taupo. P values are displayed in the upper right hand corner, F_{ST} comparisons are displayed in the lower left hand corner. | 61 |
| 4.5. Summary of stepwise backward elimination of factors model selection for 345 bp allele frequency in Lake Taupo rainbow trout. Final model and goodness-of-fit are reported | 62 |

| | |
|--|-----|
| 5.1. Microsatellite loci, fluorescent dye label used for visualization, and PCR concentration in a 5 µl reaction for seven markers amplified in this study..... | 75 |
| 5.2 Sample information, including river, sampling year, and number of samples (N), and summary statistics, including F_{IS} , allelic richness (Ar), expected heterozygosity (H_E), and the number of private alleles (PA) | 76 |
| 5.3. Population genetics statistics in scale samples compared to recent samples, including F_{IS} , allelic richness (Ar), expected heterozygosity (H_E), the number of private alleles (PA), and the total number of alleles (N_A) | 78 |
| 5.4. Pairwise estimates of genetic differentiation (F_{ST}) from five microsatellite loci. F_{ST} values are below the diagonal and P values above the diagonal, obtained after 200 permutations. The adjusted significance level after a Bonferroni correction for multiple comparisons was $P < 0.005$ | 79 |
| App. Table 3.1. Population genetic statistics for ten Lake Taupo populations and one Rotorua rainbow trout population | 100 |

Some captions are abbreviated from those used within the text.

CHAPTER 1

Introduction

1.1 Introduced populations

The success of introduced species can often be attributed to the novel environment and the conditions surrounding the founder event. If only a few individuals are introduced or are introduced from a small number of source populations, then the strong genetic drift experienced during the bottleneck can reduce genetic diversity, and overwhelm the efficacy of natural selection. In order for an introduced species to successfully colonise a new environment it must be able to persist under a different set of environmental conditions, compete with existing species which might occupy the same niche, and in some cases adapt to its new habitat.

Contemporary adaptation is the change in a heritable trait within an observable timeframe (Chakraborty and Ryman 1983; Stockwell *et al.* 2003). Colonists sometimes need to adapt in order to survive and thrive in their new environments, but the reduced genetic diversity associated with a colonization bottleneck is expected to limit their evolutionary potential. Despite these limitations, introduced species often not only survive and persist in their new environments, but also expand their range, and sometimes becoming invasive and destructive in their new habitat. The success of a bottlenecked and naïve species, which represents a repeated subversion of an apparent genetic paradox, depends largely on the effective population size of the selected colonists and number of source populations, also known as propagule pressure (Reznick and Ghalambor 2001; Allendorf and Lundquist 2003). High propagule pressure from multiple sources has in some cases resulted in higher genetic diversity in the introduced population than in each individual source population; in such a case, the genetic ability of the introduced species to respond to its environment is not constrained and is possibly enhanced (Kolbe *et al.* 2004; Roman and Darling 2007).

Contemporary adaptation has occurred in a wide variety of introduced taxa worldwide. It has been suggested that multiple introductions have played a role in the success of invasive *Anolis* lizards in Florida, where a genetic analysis indicated at least eight introductions have contributed to genetic variation, which now exceeds that found in any one native population (Kolar and Lodge 2001). The rapid evolution of *Drosophila subobscura* in the New World illustrates the incredible speed at which evolution can occur (one or two decades) (Huey *et al.* 2000). The mechanisms underlying these adaptations, however, are still not well understood, and some successful invaders actually exhibit very low levels of variability (Roman and Darling 2007). An understanding of the adaptation of introduced populations first requires an understanding of their genetic composition, and therefore it is important to attempt to trace the routes and history of invasion (Estoup and Guillemaud 2010) and describe their present level of diversity and population structure.

Freshwater fish introductions

Freshwater fish translocations are one of the oldest human-mediated forms of introduction, since there is documentation of their occurrence dating from the time of the Roman empire (Copp *et al.* 2005a). However, like many introductions, freshwater fish can be catastrophic in new environments (Kolbe *et al.* 2004). Nile perch (*Lates niloticus*) were established in Lake Victoria for sport fishing through multiple introductions in the 1950's and subsequently caused the extinction of several species of native cichlids (Vitule *et al.* 2009), forever altering the environment. The introduction of mosquitofish (*Gambusia affinis*) to Australia in the 1920s for mosquito control has also led to the decline of many native fish populations (Pringle 2005). The success of mosquitofish in these novel environments is attributable to their evolutionary potential, which, combined with their short generation time and an ability to tolerate a range of physiological conditions, have made them a model organism for studies of microevolution. Recent introduction events in freshwater systems provide unique opportunities to observe contemporary evolution *in situ*.

In addition to offering new opportunities for the study of contemporary evolution, freshwater fish introductions have been an important aspect of human history. While deliberate freshwater fish introductions have a long association with human activities, and fish are usually an important contributor to local economies, the laws governing freshwater

introductions are often less stringent than for other species (Copp *et al.* 2005a; Copp *et al.* 2005b). Freshwater fish were introduced to New Zealand in the mid-colonial period, with the first being goldfish (*Carassius auratus*) in 1864, followed by perch (*Perca fluviatilis*), trouts, salmon and many other fish towards the end of the 19th century (McDowall 1991) and smelt as food for the trout and salmon in the early 20th century (McDowall 1994). Of the salmonids introduced to New Zealand, brown trout (*Salmo trutta*), Chinook (also known as Quinnet) salmon (*Oncorhynchus tshawytscha*), and rainbow trout (*Oncorhynchus mykiss*) are now most abundant and widespread (McDowall 1991). Until the 1960s and 70s, though, relatively little trout and salmon research was done (McDowall 1991). The utilization of population genetic methods can not only help elucidate patterns of genetic structure and genetic changes associated with translocations, but may also clarify this somewhat understudied area (Hänfling 2007) and provide both scientific and economic benefits.

1.2 Population structure and management

Defining the term “population” is difficult and many different definitions have been used in population biology and management. These definitions can be roughly categorized into an ecological paradigm and an evolutionary paradigm (Waples and Gaggiotti 2006). The ecological paradigm is often delineated by demographic cohesion, where a group of individuals lives in the same space at the same time (Berryman 2002). The evolutionary paradigm is characterized by reproductive isolation, so that individuals within close spatial proximity have the potential to mate with any other member of the group (Hartl and Clark 1997). Both paradigms have been criticised since neither has thus far been largely applicable to datasets from wild populations (Waples and Gaggiotti 2006). However a population is defined, though, it is important to quantify genetic differentiation between populations in order to implement biologically sustainable management (Laikre *et al.* 2005).

Population genetic variation is often spatially structured and it can be roughly categorized into three classes based upon rates of gene flow (following Laikre *et al.* 2005). (1) Distinct populations have interpopulation gene flow low enough to permit genetic divergence from closely located populations. (2) Populations that show a continuous change have larger gene flow which results in an isolation by distance (IBD) pattern. (3) Populations which have no interpopulation differentiation have extremely high gene flow which keeps all the

individuals in a geographic region genetically similar. It is important to tailor management decisions to the type of spatial population structure in a given area, and combinations of these three models might be present in any given geographic area since differentiation is often more of a continuum than a set of discrete states (Waples and Gaggiotti 2006).

Population variation can also be temporally structured, as is the case with pink salmon (*Oncorhynchus gorbuscha*), where populations are genetically differentiated between even and odd years (Groot and Margolis 1991). While life history characteristics like spawning year can create predictable temporal genetic structure, catastrophic events like landslides can also alter the genetic structure of a population (Hansen 2002). Therefore it is important, where possible, to investigate temporal variation in genetic structure through the analysis of historical samples, particularly if using analyses which assume that population differentiation does not change over time (Waples 1998).

Population management

Fisheries management is aided by the accurate definition of management units using populations genetic tools to identify genetic structure (Palsbøll *et al.* 2007). Ideally these management groups will be genetically differentiated from their neighbours, and will be created with consideration of any seasonal fluctuations in population structure. Locally adapted groups should be managed as one unit (Conover 1998); though the only negative consequence of over-splitting these groups is an increased use of management resources (Allendorf and Luikart 2007). Treating a series of smaller populations as a large one, however, could lead to a depletion of genetic variation and a loss of local adaptations, especially in situations where harvesting pressure is not evenly distributed (Stephenson 1999). The establishment of management units previously focused on rejecting or failing to reject panmixia as a way to determine whether populations were significantly differentiated (Palsbøll *et al.* 2007), since panmictic populations are assumed to lack local adaptations.

The idea that the complexity of river habitats promotes genetically and phenotypically differentiated groups within salmon species led to the stock concept in the study and management of fisheries, particularly through the work of W.E. Ricker and several symposia by the fisheries biology community (Quinn 2005). Subsequently, genetic tools were applied to stock structure analysis, mixed-stock analysis, and conservation of threatened

species/exploited species (Carvalho and Hauser 1994). The theory of population structure within fisheries was then extended by the inclusion of the concept of evolutionary significant units (ESU) (Waples 1991; Waples 1995), which was implemented to more accurately reflect the need to preserve genetic diversity and the evolutionary process within stocks. Today a combination of ESU and management unit concepts are often used to define exploited populations, with the focus on preserving genetic diversity within populations instead of on previous concepts of rejecting panmixia (Palsbøll *et al.* 2007).

1.3 Rainbow trout

Background of the species

Rainbow trout are native to North America and native populations range from northern Alaska, just below the outlet of the Yukon River, down to Rio del Presidio, Mexico (Behnke 1992). Pacific salmonids, all members of the *Oncorhynchus* genus, differentiated in the Miocene period (Waples *et al.* 2008), likely partially due to a genome duplication which caused modern day salmonids to be tetraploid (Allendorf and Thorgaard 1984). Rainbow trout probably speciated in Southern California/Northern Mexico and then spread north to the Columbia River, on the border of what are now the states of Washington and Oregon. It is postulated that this occurred between 50,000 and 32,000 years ago, during the Pleistocene (Behnke 1992). The original classification of rainbow trout was as *Salmo gairdneri*¹ and *Salmo iridues*, for Columbia River steelhead (which had fine scales) and other coastal rainbow trout (which had course scales) respectively (Smith and Stearley 1989). Fisheries workers in the 1930s and 1940s found that different incubation temperatures produced different numbers of scales and vertebrae, and thereafter all forms of rainbow trout were regarded as *Salmo gairdneri* (Behnke 1992). Following the revelation that rainbow trout

¹ There is also some question as to whether *gairdneri* should end with a single or double “i” since it is a patronym, or species named for a person. The current convention of the American Fisheries Society as of 2004 is to use a double “i” (McDowall 2005).

were more genetically similar to Pacific salmon than Atlantic salmon, their genus was revised to *Oncorhynchus mykiss* in the late 1980s (Smith and Stearley 1989). The species name “mykiss” comes from Russian, since at this time Kamchatkan trout were also found to be the same species, and now the full range of *O. mykiss* is known to encompass Russia (Smith and Stearley 1989). These revisions signalled the end of nearly a century of debate over the taxonomy of this very phenotypically plastic, and therefore difficult to classify, genus of salmon and trout.

Rainbow trout and steelhead trout are now classified as the same species, *O. mykiss*, but the confusion over their classification stems partly from their different life history types. Steelhead trout are anadromous, meaning that they migrate to sea from their natal freshwater environment as juveniles and then mature at sea, only returning to their native habitat to spawn. Rainbow trout in contrast are adfluvial and remain resident in their natal catchment (Quinn 2005). These two life history types often co-exist and are known to interbreed (Docker and Heath 2003; Olsen *et al.* 2006).

Spawning migration behaviour varies widely among the *O. mykiss* ecotypes. Steelhead trout migrate upstream in summer, fall or winter depending on the population, but are likely to be migrating into freshwater somewhere in their native range at every month of the year (Behnke 1992). Steelhead may spawn more than once, but survival to second spawning is low (Behnke 1992). Resident rainbow trout populations often spawn in the spring, but spawning can occur from December through January and from May through June throughout their native range (Behnke 1992). This wide range in spawning time is thought to be due to rainbow trout’s long history of selective breeding in hatcheries, which eliminated the selective pressures and interspecific competition that maintained a spring spawning season (Gall and Crandell 1992).

Introduced rainbow trout

Rainbow trout are a successful introduced species on every continent except Antarctica. Introductions since 1874 have included eastern North America, Africa, Asia, Australasia, Europe, and South America (MacCrimmon 1971). Acclimatization societies around the world began culturing and shipping rainbow trout in the late 1800’s as it was a valued sport fish, could withstand warmer temperatures and more difficult conditions than

other trout, and grew more rapidly in hatcheries (Halverson 2010). There are now at least nine domesticated strains of rainbow trout reared in hatcheries in California alone (Busack and Gall 1980). Steelhead and rainbow trout are threatened and endangered in their native range (Northwest Regional Office 2011). They are highly successful, however, as an introduced species, and in their new habitats they can have a negative impact as disease vectors and aggressive predators of native fishes (Crowl *et al.* 1992). Because of its global distribution, economic importance and long history of exploitation and management, rainbow trout will likely remain one of the most heavily managed of freshwater fishes.

The genetic changes associated with introductions

Rapid evolution has been documented in introduced salmon and trout populations worldwide. Sockeye salmon (*Oncorhynchus nerka*) introduced to Lake Washington from Baker Lake in Washington state developed two different ecotypes, which evolved reproductive isolation after fewer than 13 generations (Hendry *et al.* 2000). In Patagonia, rainbow trout introduced from California around the turn of the 20th century have developed anadromy, longer life spans, and a higher rate of iteroparity, in which multiple spawnings occur throughout the life of the fish, though a genetic basis for these changes has not been investigated (Pascual *et al.* 2001). It is in fact thought that the development of anadromy is more likely a re-development, since these fish were probably sourced from anadromous populations (Behnke 2002). Chinook salmon were introduced to the South Island of New Zealand at approximately the same time as rainbow trout were introduced to Patagonia (~27 generations ago), and they have developed a distinct population structure between drainages (Hurlbert *et al.* 1972), suggesting that significant genetic changes can occur in relatively few generations.

1.4 The Lake Taupo system

Management of trout in New Zealand and Lake Taupo

There are no known populations of New Zealand rainbow trout which are anadromous; all populations are either solely riverine or adfluvial-lacustrine (McDowall 1994). This is in spite of the fact that they were likely introduced from anadromous populations in California (Scott *et al.* 1978). Rainbow trout were rapidly established in the

inland lakes in the Central North Island and the lakes along the eastern edges of the Southern Alps on the South Island (McDowall 1994). Despite the importance of recreational fishing throughout New Zealand, scientific fisheries research was only a minor component of initial management programs until the 1960s and 70s (McDowall 1991).

The Lake Taupo trout fishery is notable for large catches of brown and rainbow trout, and unusually this fishery is managed by the Department of Conservation, since Ngāti Tuwaretoa gifted the lake bed to the Crown in 1926 in exchange for annual royalties (Scrimgeour and Oxley 2001). Until a decline in stocks in the 1980s, the fishery was managed ad hoc. Subsequently, research was undertaken to attempt to gain a better understanding of trout production in the lake. The results led to the introduction of a lower bag limit, still in place today (McDowall 1991). Research and monitoring has played more of a role in the management of the fishery in the last 20 years.

Observed change in spawning time and body size

The exact characteristics of the founding stock or stocks of New Zealand rainbow trout are unknown, but populations in Lake Taupo now differ from each other in morphometric characteristics, including dissimilar parr marks, developmental characters, agonistic behaviours and adult morphometrics (Rosenau 1991). Recently Taupo trout have exhibited reduced fork length and are spawning in more concentrated runs early in the year as opposed to year-round (Maclean and Dedual pers. comm.). This intraspecific phenotypic variation might be due to adaptive differences to the environment (Beacham 1990), especially since different populations throughout Taupo can have very different spawning environments (Maclean and Dedual pers. comm.) and juvenile trout show different fat content between streams (Dedual 2002), indicating that some spawning habitats may have preferential rearing conditions. Upriver migration of salmonids varies year-to-year based on environmental conditions including temperature, river condition and flow (Quinn 2005), and therefore the reported interannual changes could be a response to environment (phenotypic plasticity) or due to selection pressures within the fishery.

1.5 Circadian rhythms and the genetic basis of spawn timing

Circadian rhythms

The circadian clock is one of the most widely studied examples of a complex phenotypic trait that is controlled by a gene-regulatory network (Kyriacou *et al.* 2008). Circadian clock gene regulatory networks are found in a wide range of taxa, from fungi, plants and insects to vertebrates (Bell-Pedersen *et al.* 2005; Yu and Hardin 2006). Though the mechanics of these networks vary between some of these taxa, all contain an autonomous circadian oscillator which involves clock-controlled gene expression (Bell-Pedersen *et al.* 2005). The *Clock* gene in particular is a transcription factor within the circadian gene regulatory network whose function was first determined in *Drosophila* (Darlington *et al.* 1998). The circadian oscillator is regulated by control of the *Clock* gene, so *Clock* is a basic helix-loop-helix transcription factor that activates the transcription of the genes *period* and *timeless* into mRNAs which are subsequently translated into proteins, but these mRNAs also abrogate the activity of *Clock* in a rhythmic fashion (Darlington *et al.* 1998). The degradation of the *timeless* protein is light-stimulated so that the *Drosophila* circadian clock is attuned to light-cycling within the environment, although in the absence of environmental cues the oscillator persists in maintaining circadian rhythms (Yu and Hardin 2006). In vertebrates the *Clock* gene includes a poly-glutamine repeat known as a poly-Q region which acts as a transcriptional trans-activation domain that influences behaviour and physiology (Gekakis *et al.* 1998).

It has recently been observed that there are latitudinal clines in variation of the poly-Q region of the *Clock* gene variation, with larger alleles usually occurring at higher latitudes (Johnsen *et al.* 2007; O'Malley and Banks 2008a). Because *Clock* is regulated by daylength, it also shows seasonal fluctuations in expression which can be attributed to daylength variation (Davie *et al.* 2009); although the effect of other environmental changes through the year, such as temperature, cannot be ruled out as a cause for the fluctuations (Tournier *et al.* 2007). This seasonal difference in length of the poly-Q region in *Clock* has been associated with egg-laying time and fledgling success (ergo fitness) in blue tits (*Cyanistes caeruleus*), suggesting that it is likely under positive selection (Liedvogel *et al.* 2009). The importance of these biological activities for fitness and survival means that there is probably intense selection on

the circadian rhythm gene regulatory network, particularly the poly-Q region, and that observed variation in that network is related to biologically important traits.

The genetic basis of spawn timing

Spawning time is controlled by multiple genes and timing is associated with genetic variation: individuals spawning at the beginning and end of the season have reduced heterozygosity (Leary *et al.* 1989). *Clock* genes are involved in both the reproductive development and timing of spawning in rainbow trout, and therefore they can influence life history differences across the lifespan of trout populations, including migration to sea as juveniles and spawning migration time as adults. Reproductive development in rainbow trout is directly related to photosensitivity and circadian rhythm (Duston and Bromage 1988), and photoperiod is used in aquaculture to control and often accelerate the maturation of farmed fish (Bromage *et al.* 2001). Since photoperiod is important to maturation, mutations in the timing genes could result not only in changes to adult spawning but also in development.

A single autotetraploidization event 25 to 100 MYA likely led to the evolution of salmonid fishes, including rainbow trout (Allendorf and Thorgaard 1984). As a result there are functional molecules, including several hormones, known to be encoded by duplicated genes in *Oncorhynchus* species (Allendorf and Waples 1996). The duplication of genes throughout the genome is an important consideration not only for functional genes, but also when interpreting non-functional microsatellite markers, especially if they are used to estimate effective population size and genetic diversity of managed populations. Duplication is also known to have occurred in salmonid *Clock* genes: *OtsClock1a* and *OtsClock1b* are two homologous copies of the *Clock* gene which have both remained functional in the salmonid genome (O'Malley and Banks 2008b).

The poly-Q region, a region which influences behaviour and physiology as described previously, has shown variability across salmonid species, particularly in the *Clock1b* gene (Paibomesai *et al.* 2010). This most likely indicates that it is an important region for altering circadian phenotypes. Changes to Q-rich regions in other species has shown to alter circadian rhythms, with deletion of amino acids in a Q-rich region in mice resulting in a lengthening of circadian period up to 4 hours (Vitaterna *et al.* 2006). Longer variants of the poly-Q region in Chinook salmon have been correlated with greater latitude, perhaps causing fish further north

to spawn later in the year when conditions are more appropriate (O'Malley and Banks 2008a). This emphasizes the importance of this region for the regulation of circadian behaviours, and particularly the potential for the regulation of spawning behaviours.

1.6 Aims of the thesis research

The overall goal of the research reported in this dissertation was to investigate the population genetic structure and spawning time genetics of Lake Taupo rainbow trout. This goal was comprised of four specific aims:

1) Tracing the origins

The aim of this section of the study was to determine the source population(s) for the New Zealand trout introduction. The origins of rainbow trout from Lake Taupo, New Zealand are investigated by comparing their genotypes at 15 microsatellite and 93 single nucleotide polymorphism (SNP) loci with those from native Californian rainbow trout and steelhead sampled from multiple geographic sites. The Californian sites were chosen as the possible source populations for the New Zealand introduction. The expectation is that Lake Taupo rainbow trout will have similar genetic characteristics to one or more of the sites of origin specified in historical records, and that those characteristics and their history of introduction will have contributed to their current population genetic structure.

2) Spatial population genetic structure

The aim of this section of the study is to determine whether there is genetic differentiation among populations sampled from tributaries around Lake Taupo. It is important to understand whether the spawning populations of different tributaries are genetically differentiated from each other, to understand both the population genetics of introduced species and, more practically, to be able to manage fisheries. Microsatellite DNA markers are used to determine the levels of genetic variation within rivers and between rivers in the Lake Taupo catchment. The null hypothesis is that there will be genetic panmixia throughout the catchment (i.e., no statistical differences will be observed in population genetic variation based on geography, lake vs. tributaries; among tributaries).

3) *Spawn timing genetics*

The aim of this section of the study is to determine whether there is a correlation between variation at the *OtsClock1b* gene and the spawn timing of rainbow trout. Recently Taupo rainbow trout showed a reduction in fork and spawning runs were more concentrated early in the year as opposed to year-round (Maclean and Dedual pers. comm.). This study investigated whether there are polymorphisms in the poly-Q area of the *OtsClock1b* gene and, if so, whether the variation is attributable to spatial structuring. Although previous studies in North America have only ever detected one allele at the *OtsClock1b* locus in rainbow trout, there could be more than one allele in the Lake Taupo populations and spatial structuring associated with allele size, as is found in Chinook salmon (O'Malley and Banks 2008a). In addition to investigating the possibility of a spatial cline in allele frequencies, seasonal structuring of *OtsClock1b* allele frequencies is only investigated, because the observed shift in spawning times may be due to a selection event within the Taupo catchment.

4) *Temporal genetic variation*

The aim of this section of the study is to determine whether there is temporal population genetic structure of rainbow trout populations in Lake Taupo. Variation at a candidate gene for selection, *OtsClock1b*, is also investigated to determine whether there is change over time at this locus. The null hypothesis is that there is no change over time, in either genetic diversity, genetic structure, or in the spawning time locus *OtsClock1b*.

CHAPTER 2

The origins of New Zealand rainbow trout (*Oncorhynchus mykiss*)

2.1 Introduction

Aquatic species have a long history of anthropogenic translocations, both intentional and unintentional (Roman and Darling 2007). Understanding the pathways and timing of introductions can help explain contemporary patterns of genetic variation. However, as the time following introduction increases, it becomes more difficult to trace the source population(s) accurately because the processes of genetic drift, mutation and natural selection will change the composition of alleles in the newly founded and source populations. Furthermore, if the size of the introduced population fluctuates below the original number of founders, it is difficult to accurately estimate the actual size or composition of the founding population (Anderson and Slatkin 2007). Bottlenecks may also lead to a significant loss of genetic diversity and a shift in allele frequencies from those found in the source populations (Frankham 2005). In addition, the difficulty of identifying the source of an introduction is often compounded by the lack of or the inaccuracy of historical documentation, which could verify potential sources.

Previous studies in a variety of taxa that have tackled this problem relied on one or two types of genetic markers to track poorly documented introductions back to their source populations (Astaneï *et al.* 2005; Kawamura *et al.* 2006). Traditionally, markers like allozymes, microsatellites, and mitochondrial DNA have been used to determine the source population of introduced freshwater species (Astaneï *et al.* 2005; Colautti *et al.* 2005), including several fish species (Quinn *et al.* 1996; Riva Rossi *et al.* 2004; Kawamura *et al.* 2006; Brown and Stepien 2009). However, these studies were often inconclusive. For instance, Riva Rossi *et al.* (2004) concluded from their mtDNA analyses that the single source of Argentinean rainbow trout was the McCloud River, but also noted that there were possible multi-source scenarios that could have led to the same outcome. Similarly, Quinn *et al.* (1996) found that although allozymes confirmed a

Sacramento River population as a source, neither allozymes nor mtDNA provided a definitive answer for the ancestral population of New Zealand Chinook salmon (*Oncorhynchus tshawytscha*) populations, most likely because of possible genetic changes over the years in both introduced stocks and Californian populations. While these studies have contributed valuable information about the genetic effects of introduction, the increased availability of high resolution molecular genetic markers like single nucleotide polymorphisms (SNPs) for non-model species offer the opportunity to tackle these problems in even finer detail, as demonstrated in marine fisheries genetics (Hauser and Carvalho 2008) and salmonids (Abadia-Cardoso *et al.* 2011; Clemento *et al.* 2011). Among molecular markers currently available, microsatellites exhibit the highest levels of polymorphism, while multilocus SNP panels can be used to genotype larger numbers of loci with less laboratory effort (Seeb *et al.* 2011). The combination of a high rate of polymorphism and a large number of loci increases the overall number of alleles. This, in turn, increases the statistical resolution when trying to identify diagnostic genotypes among populations (e.g. population genetic assignments; Narum *et al.* 2008); and when trying to correctly identify the sources of recently founded populations (Astaneï *et al.* 2005; Brown and Stepien 2009).

Rainbow trout (*Oncorhynchus mykiss*) are an excellent species to assess whether the power of larger multiple-marker datasets allows a conclusive understanding of historical translocations. They are native to the North Pacific Rim from Mexico to north-eastern Asia and have been deliberately introduced all around the world (Behnke 1992). Most introductions originated from California, particularly from the Baird station on the McCloud River (Dollar and Katz 1964; see Fig. 2.1 for location), although this was not the first or only source (Behnke 1992). There is also extensive genetic information available for rainbow trout because of their commercial and recreational importance (Rexroad *et al.* 2008).

Introduced rainbow trout have successfully colonised many river systems in New Zealand, particularly those associated with lakes (McDowall 1994). The precise origins of New Zealand rainbow trout, however, are unclear, with several authors indicating conflicting sources and dates of shipment. For example, there is inconsistency among different records regarding the shipment of rainbow trout from San Francisco to the Auckland Acclimatisation Society in New Zealand in 1883 (Scott *et al.* 1978). Hobbs

(1948) cites a memorandum from the Marine Department which indicates that trout eggs shipped to New Zealand came from the Russian River, while Stokell (1955) firmly believed that one source was Lake Almanor (Feather River, California basin). However, Scott *et al.* (1978) cite evidence which supported the hypothesis that the first introduction in 1883 came from a hatchery on Sonoma Creek, a tributary of San Francisco Bay, because no Russian River hatchery was operational before the late 1800s, and the town of Ukiah, where the hatchery was eventually located, was only accessible by stagecoach before 1887. Furthermore, although many international rainbow trout introductions came from the Baird Station on the McCloud River, Scott *et al.* (1978) could find no direct record in the U.S.A. to corroborate a shipment of trout eggs from the McCloud River to New Zealand. There are indirect mentions in several reports, but there is no record of shipment from Lake Almanor until 1930. While there was a record in Auckland of a shipment in 1878 taken from Lake Tahoe, this has since been dismissed (Scott *et al.* 1978) as Lahontan cutthroat trout (*Oncorhynchus clarki henshawi*). It is therefore possible that the Lake Taupo stock originated from a coastal Sonoma Creek population and/or the McCloud River in the late 1800s, with a later introduction from Lake Almanor in the 1930s.

Here the origins of rainbow trout from Lake Taupo, New Zealand are investigated by comparing their multilocus genotypes at 15 microsatellites and 93 SNPs with those from native Californian rainbow trout and steelhead (the anadromous ecotype of *O. mykiss*) sampled from multiple geographic sites. These Californian sites are evaluated as possible source populations for the New Zealand introduction. This study's aims are to: 1) assess the ability of SNPs and microsatellites to detect the sources of an introduction and 2) use data collected from these two marker types to determine the source populations for the New Zealand trout introduction. Our null hypothesis is that there is a single McCloud River origin. We expect that Lake Taupo rainbow trout will have similar genetic characteristics to one or more of the sites of origin specified in historical records, and that those characteristics and their history of introduction will have contributed to their current population genetic structure.

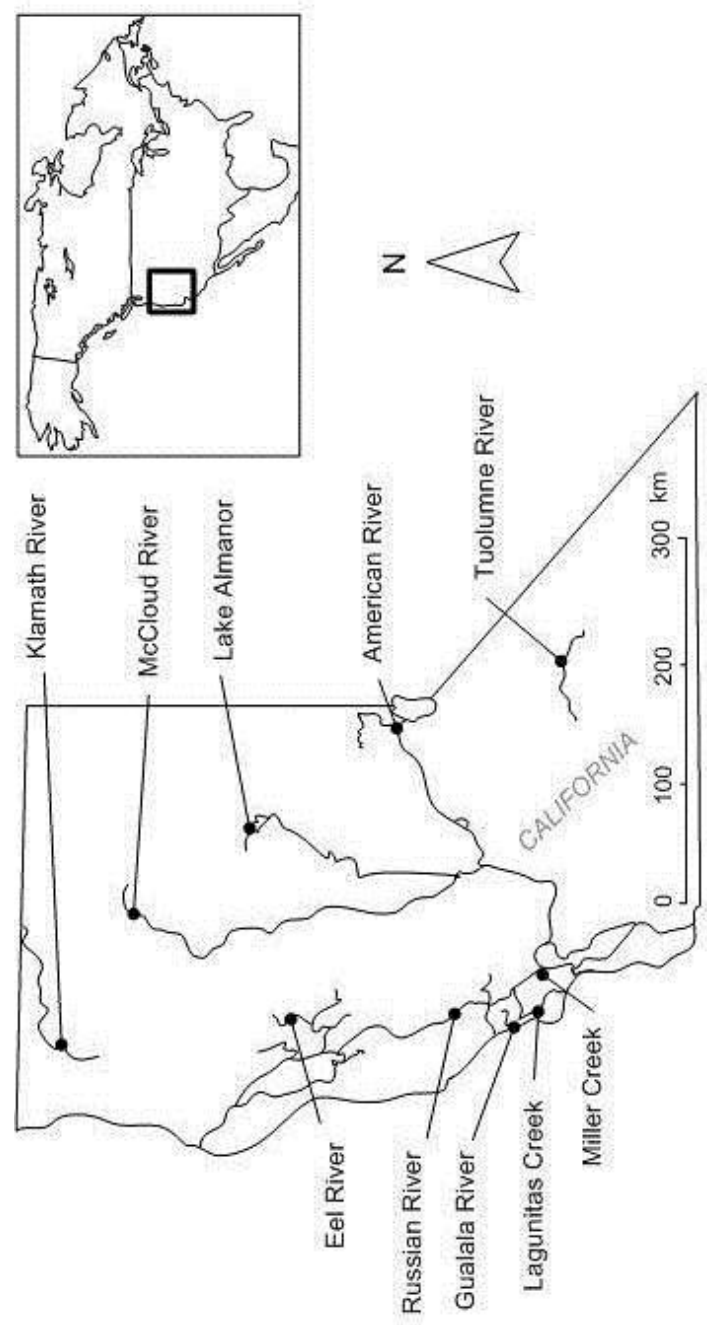


Figure 2.1. Map of California showing sample sites.

2.2 Materials and Methods

Sampling

Rainbow trout were captured at six locations throughout the Lake Taupo catchment in New Zealand from August 2006 to June 2007 at stream traps, by angling, or with electrofishing. Fifty fin clips were taken and stored in 95% ethanol at 4°C until DNA extraction.

Samples from ten Californian *O. mykiss* populations were collected between 2001 and 2006 using back-pack electro-fishers and a stratified stream sampling design intended to minimize sampling of siblings. These populations include those that were reported as sources for the Lake Taupo population (Scott *et al.* 1978); the Upper Feather River sample was taken just downstream from the outflow of Lake Almanor and Miller Creek was used as a proxy for Sonoma Creek, since their outflow into San Francisco Bay are only ~12km apart and Sonoma Creek is now highly degraded (Figure 2.1).

DNA Extraction and Genetic Analyses

Genomic DNA was extracted from the Lake Taupo samples using a proteinase K digestion followed by salt extraction. The remaining pellet was subjected to two ethanol washes and resuspended in TE buffer. Total DNA was quantified using a spectrophotometer (NanoDrop™ ND-1000, Thermo Scientific). Genomic DNA was extracted from Californian samples using QIAGEN DNeasy kits on a QIAGEN Bio Robot 3000 (Qiagen Inc.).

Samples were genotyped using 15 microsatellite loci following polymerase chain reaction (PCR): *Omy1011*, *Omy77* (Morris *et al.* 1996), *One11b*, *One13b* (Scribner *et al.* 1996), *Ots103*, *Oki23* (Smith *et al.* 1998), *OtsG243*, *OtsG249b*, *OtsG409*, *OtsG43*, *OtsG85*, *OtsG253b*, *OtsG3* (Williamson *et al.* 2002), *Ots1b* (Banks *et al.* 1999), and *Ssa289* (McConnell *et al.* 1995). Thermal cycling was performed on MJ Research Thermocyclers (PCR conditions are available from the authors upon request). Amplified PCR products were analysed on ABI 377 automated DNA sequencers (Applied Biosystems). All genotypes were called and confirmed by two people independently.

Ninety-three SNP loci including *Omy_AldA*, *OMGH1PROM1-SNP1*, *Omy_aspAT-123*, *Omy_COX1-221*, *Omy_nramp-146*, *Omy_Ogo4-304*, *OMY_PEPA-INT6*, *ONMYCRBF_*

1-SNP1 (Aguilar and Garza 2008), *Omy_arp-630*, *Omy_g12-82*, *Omy_gh-475*, *Omy_gsdf-291*, *Omy_mapK3-103*, *Omy_mcsf-371* (Campbell *et al.* 2009), and eighty-five loci from Abadía-Cardoso *et al.* (2011) were then genotyped. These SNPs were discovered through sequencing of rainbow trout expressed sequence tags (ESTs) in public databases. See Abadía-Cardoso *et al.* (2011) for details. A PCR pre-amplification was first performed, in order to increase the DNA fragments of interest, with the following reagent concentrations: 2.5 μ L of 2X Master Mix (QIAGEN Inc.), 1.3 μ L of pooled 0.2X TaqMan (5' nuclease allelic discrimination) assays (Applied Biosystems Inc.) and 1.6 μ L of unquantified template DNA diluted 1:2 in ddH₂O. Pre-amplification thermal cycling consisted of an initial denaturation of 15 min at 95°C, and 13 cycles of 15s at 95°C, 4 min at 60°C for (+1°C/cycle). Pre-amplification PCR products were diluted 1:3 in 2 mM Tris. High-throughput genotyping was then performed on 96.96 Dynamic Arrays (Fluidigm Corporation), which use nanofluidic circuitry to simultaneously genotype up to 96 samples with 96 assays. The preamplified samples were mixed with 1.25 μ L of 40X TaqMan SNP assay (Applied Biosystems), 2.5 μ L of 2X Assay loading reagent (Fluidigm Corporation), 0.25 μ L of 50X ROX (Fluidigm Corporation) and 1 μ L of ddH₂O. The 10x SNP assays preparation consisted of 2.1 μ L of pre-amplified DNA, 2.5 μ L of TaqMan Universal Master Mix (Applied Biosystems Inc.), 0.25 μ L of 20X GT Sample loading reagent (Fluidigm Corporation), 0.05 μ L of AmpliTaq Gold DNA polymerase (Applied Biosystems Inc.), and 0.10 μ L of ddH₂O. All thermal cycling and imaging was performed per the manufacturers' recommendations. Genotypes were then called using the Fluidigm SNP Genotyping Analysis Software.

Data Analysis

Lake Taupo samples were treated as a single population because previous analyses showed no significant genetic differentiation between sites (ERH unpublished data). Fifty Lake Taupo individuals were randomly chosen to represent the combined, introduced population. Departures from Hardy-Weinberg (HWE) and gametic phase equilibria were examined using GENEPOP on the web (Raymond and Rousset 1995). Comparisons were considered significant if $P < 0.05$ after a sequential Bonferroni correction (Holm 1979). Pairwise F_{ST} and expected heterozygosity were calculated using FSTAT 2.9.3 (Goudet 1995). Allelic richness was calculated using HP-RARE V1.0 (Kalinowski 2005), which uses a rarefaction method to account for different sample sizes. An exact binomial test was used to

compare differences in expected heterozygosity and allelic richness between Californian and Lake Taupo populations.

The distribution of allele frequencies for all microsatellite markers in the Californian populations and Lake Taupo was graphically represented, as this can be useful to detect bottlenecks by evaluating whether there has been a reduction in the proportion of rare alleles in the founded population (Luikart *et al.* 1998b). In addition, assuming a single-source scenario in which all of the loss in heterozygosity occurred in the first post-bottleneck generation (Kalinowski *et al.* 2010), the effective number of founders was estimated by comparing mean H_E in putative source populations and Lake Taupo with the equation: $H_{TAUPO} = (1 - 1/2N_f) H_{SOURCE}$ where N_f is the effective number of founders (Allendorf and Luikart 2007, Eq. 6.6). The estimate of N_f was used to calculate the frequency threshold at which there would be a 5% binomial probability of failing to detect alleles from the source population in Lake Taupo, assuming the loss of alleles strictly occurred during the founding event. This calculation was applied to both marker types, but for SNPs the expected heterozygosity was higher in Lake Taupo than some of the sources.

Principal components analysis (PCA) based on allele frequencies was used to examine population differentiation and was performed using MINITAB 13 (Minitab Inc., State College, PA, U.S.A.). One allele at each locus was excluded to account for non-independence among alleles within loci for microsatellites and SNPs, and the PCA was done on a covariance matrix.

Of the 93 SNPs that were analysed, many of the Californian populations had one or more SNPs which were fixed for one genotype (e.g. AA). To investigate which Californian populations might be sources for the Lake Taupo population, SNPs fixed in Californian populations were compared to data from Lake Taupo trout. The pattern of genetic differentiation between populations at SNPs and microsatellites, as measured by F_{ST} , was compared using a Mantel test in the vegan package (1994) for R 2.10.0 (R Foundation for Statistical Computing, Vienna). An F_{ST} outlier analysis of the SNP loci in Californian populations was done using BAYESCAN (Foll and Gaggiotti 2008).

Simulations

In order to compare the efficiency of microsatellite and SNP markers to accurately detect the true source of the New Zealand population, bottleneck simulations were performed under various scenarios. Allele frequencies from one Californian population (microsatellite and SNP data) were used to generate a starting population of 10 and 50 founders in BOTTLESIM 2.6. Only the results of 10 founders are shown as this is the most conservative estimate of the two, and a single source is even more easily identifiable with a 50-founder scenario. Random mating was assumed, with longevity set at five years and individuals reaching sexual maturity at two years (mark-recapture experiments during the 1980's indicate that about 70% of Taupo rainbow trout return as 3 year olds; Department of Conservation, unpublished data). The simulated population was allowed to grow every year until it reached 5000 individuals after 70 years (increasing by five individuals every year until it reached 50, then by 50 individuals every year until it reached 1000, and then by 100 individuals every year until it reached a maximum size of 5000 individuals at 70 years, and was maintained at that size for another 30 years (100 years total). This process was repeated 1000 times; for each replicate, the simulated genotype data option of BOTTLESIM was chosen and genotypes from 50 individuals were randomly selected from the resulting population. Batch files containing these data and the observed data from the source and all other California populations were then created to compare F_{ST} values between simulated data and observed populations in ARLEQUIN 3.5 (Excoffier *et al.* 2005). For simulated data, 95% confidence intervals (CIs) for mean F_{ST} values were obtained by discarding the lowest and highest 25 values (2.5%) for each set of simulations. For the observed data from the Lake Taupo and 10 Californian populations, 95% CIs were computed by bootstrapping loci 1000 times. Additional simulations were performed by using the genotype data of two (either two inland populations or one inland and one coastal population) and three different Californian populations as sources in order to evaluate introduction scenarios from multiple sources (with 20 total founders for all scenarios involving more than one source population). Simulation results for microsatellites were also recoded so that all but the most frequent allele at each locus were pooled (following McDonald 1994), to investigate whether differences in F_{ST} between SNPs and microsatellites could be attributed to differences in polymorphism. Functions to convert data to and from BOTTLESIM format, to subsample individuals within simulated datasets, to combine simulated data with empirical data, and to create ARLEQUIN batch files are available in the R package PopGenKit (Rioux Paquette 2011).

2.3 Results

Genetic variation within populations

Overall, neither the microsatellites nor SNPs had substantial significant departures from Hardy-Weinberg proportions or gametic disequilibrium. No locus had more than four significant HWE comparisons and the significant gametic disequilibrium comparisons were spread across all microsatellite and SNP loci, and populations, indicating that the significant results were likely the result of type-1 error due to the overall number of comparisons made. The Lake Taupo sample had one allele fixed at microsatellite *OtsG409* and therefore the Lake Taupo data at that locus were not included in these comparisons. The Lake Taupo population was also fixed for seven SNP markers, McCloud was fixed for six markers, Russian and Klamath for two each, and Miller and Tuolumne for one each.

An F_{ST} outlier analysis in Californian populations indicated that eight SNPs were candidates for selection, five for positive selection and three for balancing selection ($P = 0.99$).

Lake Taupo trout had significantly lower allelic richness and heterozygosity at 10 of the 15 microsatellite loci ($P < 0.05$; Table 2.1) and a lower proportion of rare alleles than McCloud River (24 %; Figure 2.2), but no mode shift was detected. At the 93 SNP loci Lake Taupo trout did not have significantly different heterozygosity, but did have more fixed loci than any Californian population (Table 2.1).

Genetic variation between populations

The Lake Taupo population was significantly differentiated from all Californian populations at all genetic loci used in this study, with a range of Lake Taupo-California pairwise F_{ST} between 0.137 - 0.188 for SNPs and microsatellites combined (Table 2.2). F_{ST} values overall were higher at SNP markers than microsatellites with a range of Lake Taupo-California pairwise F_{ST} between 0.137 - 0.211, as compared to 0.087 - 0.160 for microsatellite markers (Table 2.3). There was a strong correlation between F_{ST} values derived from both types of markers, as revealed by a Mantel test ($r = 0.556$, $P < 0.001$; Figure 2.2).

Table 2.1. Sample size (n), expected heterozygosity (H_E) and allelic richness (Ar) for Lake Taupo, New Zealand and California, USA populations computed from 15 microsatellite loci and expected heterozygosity.

| Population | microsatellites | | | SNPs | | | All | |
|------------|-----------------|-------|------|------|-------|------|-----|-------|
| | n | H_E | Ar | n | H_E | Ar | n | H_E |
| Taupo | 50 | 0.606 | 5.23 | 66 | 0.330 | 1.91 | 42 | 0.369 |
| Almanor | 53 | 0.686 | 9.34 | 52 | 0.314 | 1.97 | 52 | 0.366 |
| American | 60 | 0.702 | 8.77 | 60 | 0.380 | 2.00 | 60 | 0.425 |
| Eel | 63 | 0.703 | 8.88 | 32 | 0.336 | 1.96 | 32 | 0.387 |
| Gualala | 64 | 0.691 | 8.29 | 32 | 0.384 | 1.99 | 32 | 0.427 |
| Klamath | 71 | 0.700 | 8.88 | 72 | 0.303 | 1.96 | 72 | 0.359 |
| Lagunitas | 61 | 0.705 | 9.50 | 61 | 0.377 | 2.00 | 61 | 0.423 |
| McCloud | 54 | 0.635 | 7.15 | 38 | 0.319 | 1.92 | 38 | 0.363 |
| Miller | 69 | 0.667 | 7.33 | 62 | 0.361 | 1.98 | 62 | 0.404 |
| Russian | 62 | 0.646 | 7.00 | 64 | 0.369 | 1.97 | 64 | 0.408 |
| Tuolumne | 47 | 0.684 | 8.62 | 47 | 0.325 | 1.98 | 47 | 0.375 |
| Total | | | | | | | | |
| California | 653 | 0.682 | 11 | 586 | 0.347 | 2.00 | 562 | 0.394 |

The greatest pairwise F_{ST} values were between Lake Taupo and Klamath River (0.194), followed by Lake Taupo and Russian River (0.186), while the lowest values were between Lake Taupo and Gualala (0.138), American (0.139) and McCloud (0.141) rivers. Based on the 5% difference in H_E between the McCloud River and Lake Taupo at microsatellites, the estimated number of Lake Taupo founders was 10 individuals, which translates into a 0.05 probability of missing alleles with frequencies of 26% in McCloud; *OtsG43*, the 153 bp allele is present at 42% in McCloud and 0% in Lake Taupo (Figure 2.3). The most common allele at that locus in Lake Taupo is the 165 bp (at 53%), which is present at 23% in Lake Almanor, 3% in McCloud River, and 29% in Miller Creek (the coastal Creek). The same analysis in other putative sources also showed the same pattern of presence of common alleles in potential progenitor populations that were not present in Lake Taupo (data not shown).

Table 2.2. Pairwise F_{ST} values from comparisons of ten *O. mykiss* populations from California and Lake Taupo, calculated from 93 SNP loci and 15 microsatellite loci. All values are significantly different from zero ($P < 0.05$).

| | Taupo | American | Eel | Feather | Gualala | Klamath | Lagunitas | McCloud | Miller | Russian | Tuolomne |
|-----------|-------|----------|-------|---------|---------|---------|-----------|---------|--------|---------|----------|
| Taupo | -- | | | | | | | | | | |
| American | 0.173 | -- | | | | | | | | | |
| Eel | 0.139 | 0.110 | -- | | | | | | | | |
| Feather | 0.160 | 0.080 | 0.188 | -- | | | | | | | |
| Gualala | 0.138 | 0.081 | 0.064 | 0.151 | -- | | | | | | |
| Klamath | 0.194 | 0.113 | 0.167 | 0.118 | 0.139 | -- | | | | | |
| Lagunitas | 0.147 | 0.092 | 0.055 | 0.183 | 0.030 | 0.157 | -- | | | | |
| McCloud | 0.141 | 0.111 | 0.196 | 0.124 | 0.172 | 0.162 | 0.191 | -- | | | |
| Miller | 0.173 | 0.132 | 0.102 | 0.219 | 0.079 | 0.190 | 0.056 | 0.218 | -- | | |
| Russian | 0.186 | 0.116 | 0.093 | 0.203 | 0.073 | 0.183 | 0.059 | 0.208 | 0.086 | -- | |
| Tuolomne | 0.166 | 0.087 | 0.153 | 0.113 | 0.128 | 0.158 | 0.141 | 0.121 | 0.191 | 0.159 | -- |

Table 2.3. Comparison of pairwise F_{ST} values from comparisons of 10 California population and Lake Taupo rainbow trout. Comparisons were made using data from 15 microsatellite loci (below the diagonal) and 93 SNP loci (above the diagonal). All values are statistically greater than zero ($P < 0.05$).

| | Taupo | Almanor | American | Eel | Gualala | Klamath | Lagunitas | McCloud | Miller | Russian | Tuolumne |
|-----------|-------|---------|----------|-------|---------|---------|-----------|---------|--------|---------|----------|
| Taupo | -- | 0.189 | 0.138 | 0.159 | 0.137 | 0.211 | 0.152 | 0.150 | 0.186 | 0.187 | 0.186 |
| Almanor | 0.087 | -- | 0.087 | 0.213 | 0.176 | 0.139 | 0.214 | 0.144 | 0.255 | 0.223 | 0.127 |
| American | 0.124 | 0.056 | -- | 0.120 | 0.091 | 0.124 | 0.103 | 0.119 | 0.144 | 0.120 | 0.091 |
| Eel | 0.127 | 0.086 | 0.059 | -- | 0.067 | 0.191 | 0.053 | 0.227 | 0.108 | 0.096 | 0.166 |
| Gualala | 0.132 | 0.085 | 0.056 | 0.049 | -- | 0.163 | 0.028 | 0.193 | 0.086 | 0.071 | 0.139 |
| Klamath | 0.097 | 0.057 | 0.076 | 0.072 | 0.077 | -- | 0.185 | 0.185 | 0.226 | 0.206 | 0.178 |
| Lagunitas | 0.115 | 0.073 | 0.050 | 0.042 | 0.037 | 0.059 | -- | 0.215 | 0.059 | 0.061 | 0.156 |
| McCloud | 0.097 | 0.070 | 0.085 | 0.081 | 0.117 | 0.091 | 0.105 | -- | 0.247 | 0.222 | 0.130 |
| Miller | 0.098 | 0.089 | 0.085 | 0.065 | 0.067 | 0.065 | 0.048 | 0.100 | -- | 0.090 | 0.215 |
| Russian | 0.160 | 0.132 | 0.100 | 0.077 | 0.085 | 0.100 | 0.053 | 0.145 | 0.073 | -- | 0.168 |
| Tuolumne | 0.126 | 0.063 | 0.070 | 0.083 | 0.099 | 0.087 | 0.084 | 0.093 | 0.101 | 0.124 | -- |

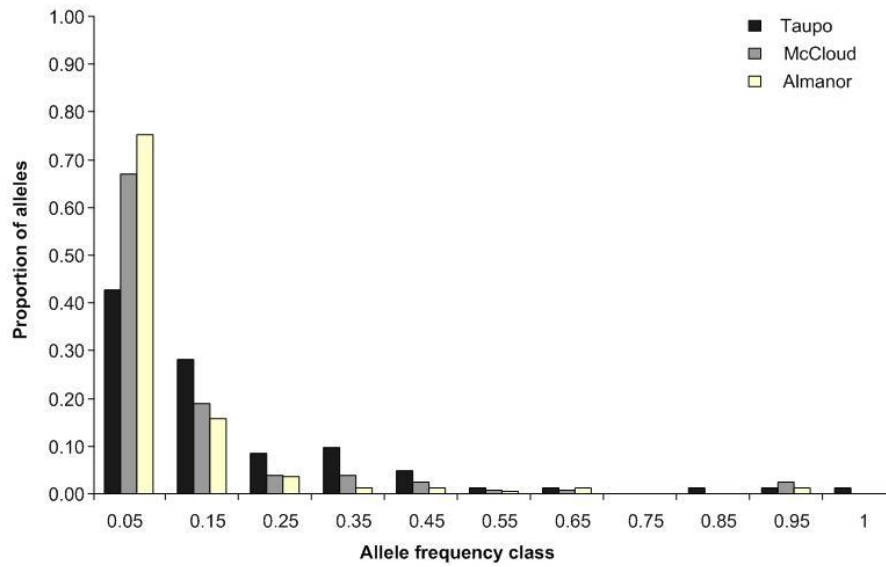


Figure 2.2. Allele frequency distribution for Lake Taupo, Lake Almanor, and the McCloud River rainbow trout populations, calculated from 15 microsatellite loci. The number under each set of bars is the median of that allele frequency class, with the number of fixed alleles shown in class “1”.

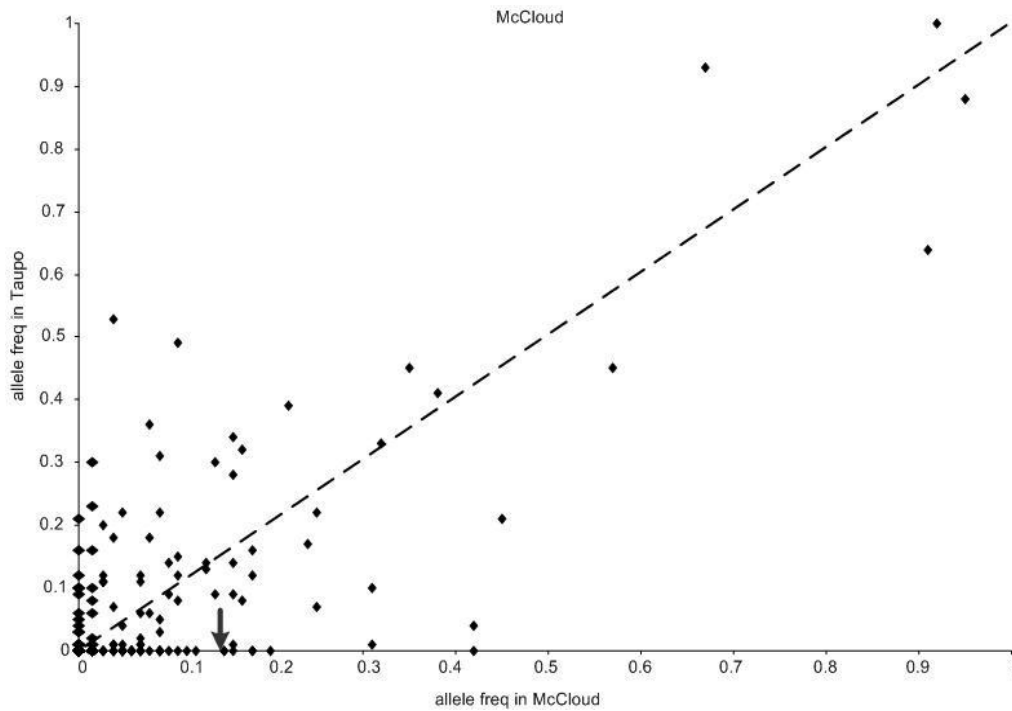


Figure 2.3. A plot of the allele frequencies for all alleles in the Lake Taupo and McCloud River populations at 15 microsatellite loci. If McCloud is the sole source, there is less than a 5% chance of missing an allele present at 14% in the McCloud River during the founding event (denoted by the arrow on the x axis).

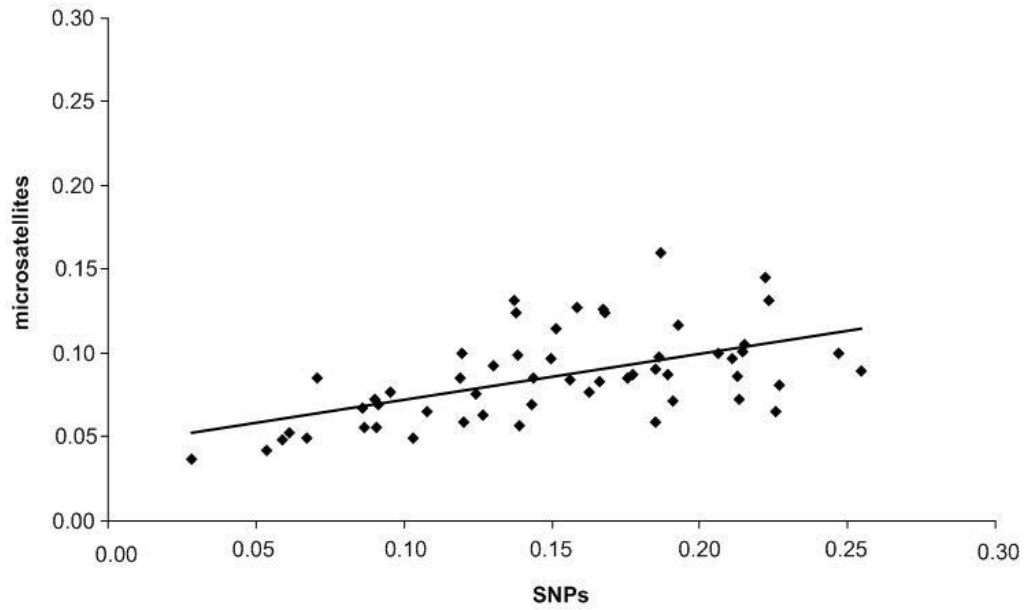


Figure 2.4. A plot of microsatellite and SNP pairwise F_{ST} values for the Lake Taupo and ten Californian populations.

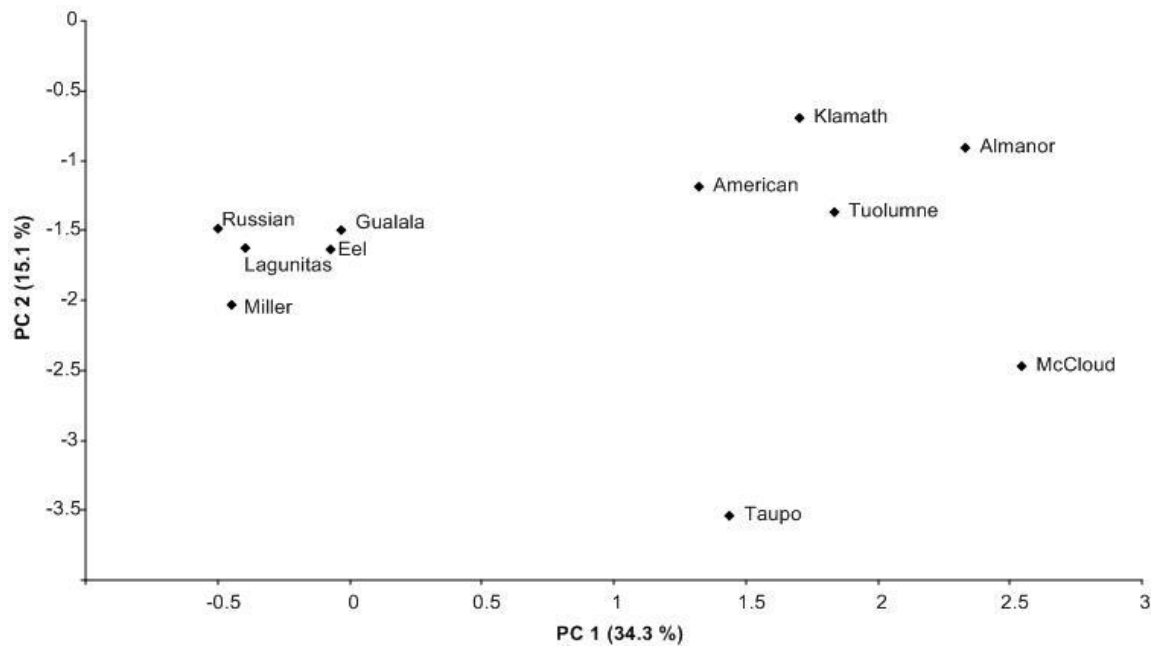


Figure 2.5. A principal component analysis (PCA) computed from 15 microsatellite loci and 93 SNP loci for Lake Taupo and ten Californian populations.

In all ten Californian populations, except Lagunitas Creek and American River, there was at least one of the 93 SNPs that was fixed in that Californian population but not fixed in the Lake Taupo population. A PCA of both the combined SNP and microsatellite data sets showed differentiation along PC axis 1 of coastal and inland Californian populations, with Lake Taupo most closely clustering with inland populations (Figure 2.5).

Simulations

Bottleneck simulations from the empirical data suggest that even with only 10 founders, both the microsatellite and the SNP markers would reliably detect the real source if the introduction was from a single source, regardless of the population chosen as the actual source (Figure 2.6a). Therefore, only results from simulations carried out with McCloud are presented. For both types of markers, the real source had the lowest F_{ST} in all replicates, and the mean difference in F_{ST} between the real source and the second closest population was 0.114 for SNPs (95% CI = 0.078 – 0.155) and 0.085 for microsatellites (95% CI = 0.047 – 0.135). In the observed results (for microsatellites and SNPs, Table 2.2 and Figure 2.6e), the difference in F_{ST} between the Lake Taupo and nine of the 10 Californian populations was smaller than the lower boundary of the 95% CI calculated from single-source simulations, suggesting that more than a single population contributed to the Lake Taupo population. For the scenario involving two inland populations (McCloud River and Lake Almanor; Figure 2.6b), 98.4% (microsatellites) and 94.4% (SNPs) of replicates correctly identified the two sources with the two lowest F_{ST} values. However, the difference between the F_{ST} values of sources and those from other populations was smaller than for single-source simulations. On the other hand, when one coastal and one inland population were involved (McCloud River and Miller Creek; Figure 2.6c), results were considerably different: the two correct sources were identified in 73.6% of replicates for microsatellites, but only 0.8% for SNPs. The coastal population of origin was mostly identified correctly (83.4% of replicates) but not the inland population of origin (5.8%). When a third coastal population was added as a source (Miller Creek, Figure 2.6d), the lowest F_{ST} values were also spread among the correct and incorrect sources. For microsatellites, the three lowest F_{ST} values were with the actual sources in 69.6% of replicates, but never for the SNP data, mostly because the American River population had the smallest F_{ST} value in 95% of replicates. Interestingly, the American River population also exhibited a lower observed F_{ST} than putative source populations (McCloud River, Lake Almanor, and Miller Creek). Correlations between F_{ST} values from

the empirical data and those from simulations were 0.27 (SNPs) and 0.68 (microsatellites) for the single-source scenario; 0.07 (SNPs) and 0.86 (microsatellites) for the scenario with two inland sources; 0.58 (SNPs) and 0.73 (microsatellites) for the scenario with one coastal and one inland source; and 0.61 (SNPs) and 0.91 (microsatellites) for the three-source scenario. Across all simulations, estimated F_{ST} values from pseudo-biallelic microsatellite data were generally intermediate between the values for SNPs and microsatellites, but the variance in F_{ST} was much greater, as illustrated by wider CIs, even for the observed results (Fig 2.6e). Finally, additional simulations with 50 founders showed that the difference between F_{ST} values from real sources and those from other Californian populations increased with the number of founders (results not shown).

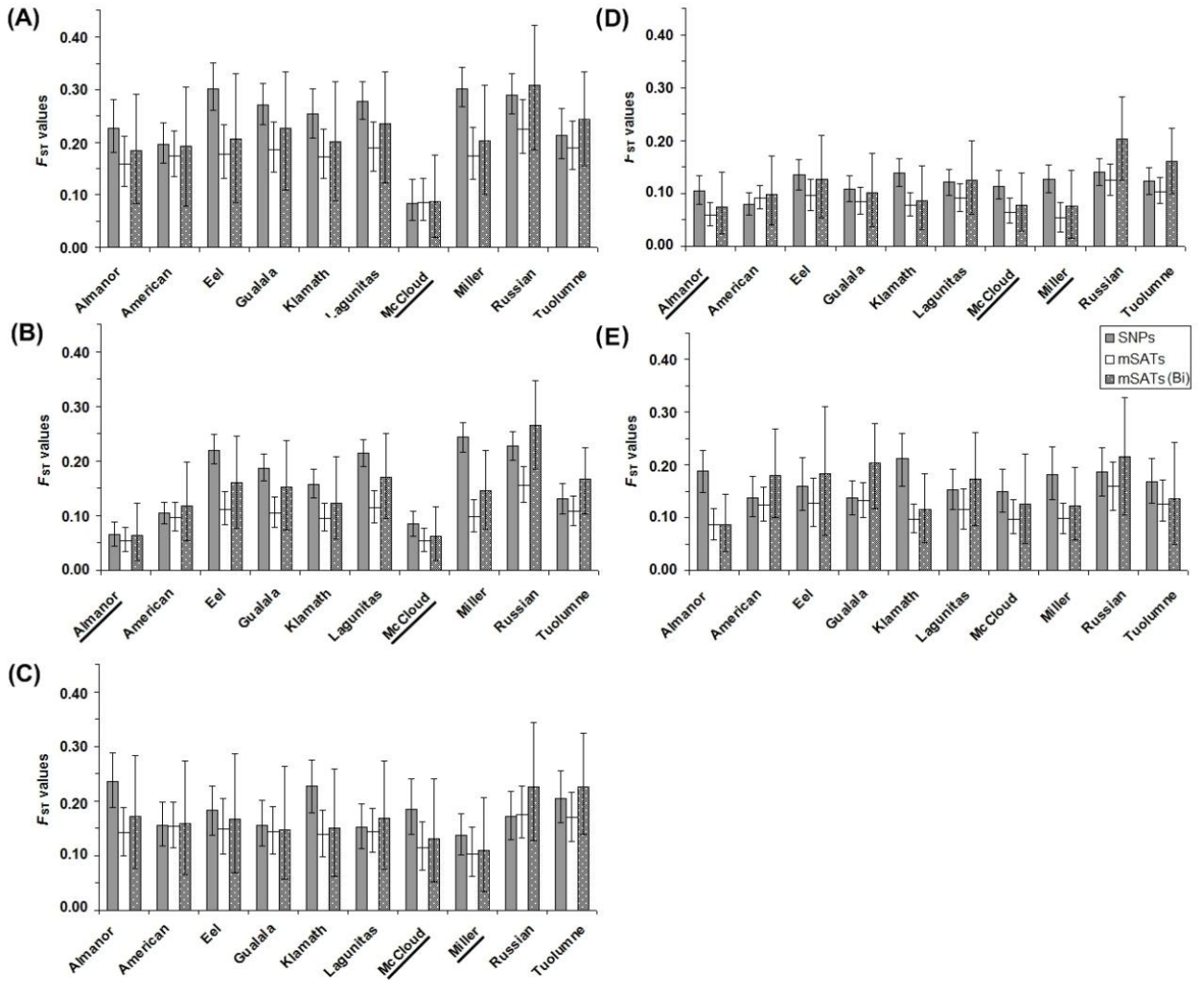


Figure 2.6. Mean F_{ST} values calculated between 1000 simulated introduced populations and ten Californian rainbow trout when one population (A), two inland populations (B), one coastal and one inland populations (C), and three Californian populations (D) contributed to the introduction, in comparison with the empirical data (E). For A-D, the populations used as sources at the start of simulations are underlined. Results are indicated for SNP markers (SNPs), microsatellite markers (mSATs) and microsatellite markers recoded as biallelic markers (mSATs(Bi)). Error bars for simulations represent 95% confidence intervals obtained by eliminating the top and bottom 2.5% replicates for each population. For the empirical data, 95% CIs were computed by bootstrapping loci 1000 times.

2.4 Discussion

What pattern of introduction do the data support?

The results from both microsatellite and SNP data strongly suggest that Lake Taupo has experienced multiple rainbow trout introductions from more than one source, so the null

hypothesis is rejected, that the McCloud River was the single source. Historical records have suggested that the McCloud River, Lake Almanor, the Russian River, and a San Francisco Bay creek are all the possible source populations for the Lake Taupo introduction. Based upon microsatellite allele frequencies, Lake Almanor and McCloud River are the most similar to Lake Taupo, and Russian River the most differentiated of all the Californian population used for comparison. There are, however, some common McCloud River alleles, which are not present in Lake Taupo trout, one of which is statistically significant. This absence of common alleles is indicative of multiple introductions, since one would expect an introduced population to contain the most common alleles from its source. At the SNP markers, only two California populations, Lagunitas Creek and American River, were polymorphic for all markers that were also polymorphic in the Lake Taupo population. Therefore these would be identified as source populations if Lake Taupo only contains polymorphisms introduced from source populations and there have been no recent mutations, and if there have been no alleles in the source populations lost either because of drift or selection, both of which are unlikely to be true. These populations, however, are not sources reported in historical records. The historical records also argue against a single source, and are consistent with the genetic evidence that multiple introductions have contributed to a mixed stock in Lake Taupo. The Lake Taupo population also has lower expected heterozygosity at the SNP markers than several of the putative source populations, which one would not expect after an introduction founder event; this also likely reflects a mixed origin. Overall it can be concluded from the combined evidence that multiple source populations likely contributed to the rainbow trout population in Lake Taupo, New Zealand.

Although Scott *et al.* (1978) suggest that New Zealand rainbow trout originated from a Sonoma Creek coastal population, there is little genetic evidence to corroborate their assertion. Acclimatisation records show that multiple source populations contributed to the Lake Taupo founding, but it is possible that the translocations from particular sources, like the McCloud River, made the greatest genetic contribution to the introduction. A PCA analysis of data from both marker types showed that Lake Taupo most closely clustered with inland Californian populations rather than coastal populations. Our simulations showed that, had there been a single source, then the expectation would be to find a significantly lower F_{ST} value when comparing that source to the Lake Taupo population than for any pairwise comparison to another Californian population. This pattern was not found with either the observed SNP or microsatellite data, as multiple Californian populations had similar F_{ST}

values to Lake Taupo. Alternatively, the most common alleles in these populations could have become more prominent in Lake Taupo through strong drift or selection. Because there were likely multiple introductions, the data are inconclusive in determining which populations were sources for the Lake Taupo introductions.

Previous studies showed that introduced rainbow trout tend to lose some rare alleles, but maintain much of their original genetic diversity (Quinn *et al.* 1996). Lake Taupo rainbow trout, however, have significantly lower allelic richness and heterozygosity, proportionally fewer rare alleles, and more fixed SNP loci than Californian populations. Bottlenecks are known to substantially reduce genetic diversity and change allele frequency distributions (England *et al.* 2003; Frankham 2005; Smith *et al.* 2005). In some cases the loss of genetic diversity can be counteracted by high propagule pressure, as happens when there are multiple introductions or a large number of founding individuals (Allendorf and Lundquist 2003; Roman and Darling 2007). Multiple introductions from diverse sources can lead to a population that is more genetically diverse than the individual source population, as has happened in *Anolis* lizards in Florida (Kolbe *et al.* 2004). A diverse genetic background from multiple founding populations may have helped trout more successfully adapt to their new habitat in Lake Taupo because of greater genetic diversity than was contained within their individual founding populations (Ellstrand and Schierenbeck 2000; Mooney and Cleland 2001). Nevertheless, New Zealand trout still show signs of a founder effect, as do introduced Chinook salmon (Quinn *et al.* 1996). This indicates that the propagule pressure was not high enough to counteract the loss of genetic diversity due to the founding bottleneck. Despite this, both salmonids are examples of genetically depauperate colonial populations succeeding despite the potentially deleterious effects of introduction.

What genetic patterns would be expected for different introduction scenarios?

The simulation results for both marker types suggest that a single-source scenario is unlikely for the introduction of rainbow trout into Lake Taupo, New Zealand. A clear pattern can be expected to emerge with both microsatellites and SNPs if only one population contributed to the introduction into Lake Taupo, with the source population exhibiting significantly lower differentiation for the two marker types when compared to non-source populations. When two inland sources are taken into consideration, the signal weakens, and can completely disappear when there are more than two source populations, particularly with

the addition of a coastal population, which makes discerning the true source purely from pairwise F_{ST} comparisons untenable. However, the situation is more complicated than simulations can show. While these analyses assume that the current genetic composition of Californian populations reflects their historical composition (at the time of the exports to New Zealand), hatchery stocking and anthropogenic alterations of the river drainages have modified the genetic signature of all salmonid populations in California, including the putative sources (Pearse and Garza 2008; Pearse *et al.* 2011). Finally, the only evolutionary force considered in the simulations is drift, but it is possible that selection may also affect some loci, particularly those eight SNP loci identified by the outlier analysis as candidates for selection.

Do SNP markers and microsatellites provide comparable results?

The SNP and microsatellite markers used in this study provided comparable genetic results, and both marker types provided consistent findings on the origins of Lake Taupo trout. F_{ST} values were significantly higher at SNP loci than at microsatellites, but this is at least in part attributable to differences in polymorphism. Indeed, when the microsatellite data were recoded to biallelic data (Figure 2.6), F_{ST} values were intermediate or even higher than those from SNPs. Overall, the SNP and microsatellite loci showed similar patterns of divergence between populations, which indicates that SNPs and microsatellites provide similar information, even when some of the loci do not reflect purely neutral processes, and is consistent with the findings of other comparative studies (Narum *et al.* 2008; Glover *et al.* 2010). However SNP data are more easily combined between labs than microsatellite data, and have lower genotyping error and mutation rates, which are very useful when either trying to identify the source population of an introduced species or understanding evolutionary change since introduction. For these reasons SNPs are preferable to microsatellites when choosing markers for introduction studies, although one needs to take into account the smaller mean number of alleles and consequent need for more SNP loci for similar statistical power. As SNP discovery and SNP typing methods become increasingly available to those studying non-model species, they will become the marker of choice for population genetic study of introduced and invasive species.

Conclusions

While it was not possible to definitively identify the source populations of Lake Taupo rainbow trout, the data and analyses presented here indicate that is most likely that McCloud River and Lake Almanor (Feather River) are the primary contributors, although a coastal population may have contributed as well. Our simulations show that a multiple source population scenario reduces pairwise F_{ST} values for non-source population comparisons and makes the determination of the true source difficult. We also found that SNPs and microsatellites produce comparable results. However, the ability to easily combine SNP data between labs, coupled with their greater stability and higher genotyping throughput, will probably make them a preferable genetic marker for tracing the history and spread of introductions and biological invasions.

Ultimately studies like this give insight into how species will respond to new environments by studying populations that have been introduced into novel environments. As Behnke (2002) observed, understanding the true sources is also important for understanding life history characteristics: in the case of Patagonian rainbow trout, the assumption of a McCloud River sole source could have led to the claim of re-evolved anadromy (Pascual *et al.* 2001) when, in fact, the trout were expressing a trait they inherited from another ocean-going source. By understanding the population genetic changes that many introduced populations undergo to persist in their new environments, researchers and managers can better anticipate populations' response to changing conditions and design management actions to better mitigate potential negative outcomes.

CHAPTER 3

Population genetic structure of the rainbow trout in Lake Taupo, New Zealand

3.1 Introduction

Freshwater fish populations often have high interpopulation genetic differentiation since connectivity between rivers and lakes is restricted, which limits the size of subpopulations (Ward *et al.* 1994). Due to the geographic constraints to migration, freshwater populations tend to have a lower level of overall genetic variation, evidenced by both lower heterozygosity and number of alleles per locus (DeWoody and Avise 2000; McCusker and Bentzen 2010). The lower effective population sizes of freshwater fishes relative to marine populations, which are estimated to be almost 10-fold higher, are partially attributable to this lower genetic variation (DeWoody and Avise 2000). One would expect that low genetic variation of freshwater fishes would be disadvantageous (e.g., fewer genetic potential for adaptation to new environmental conditions) when they are introduced to new environments. Approximately 63% of freshwater fish introductions worldwide are successful (Ruesink 2005), indicating that the majority of introductions are successful in spite of the limiting factors.

The genetic variation of introduced populations largely depends on the conditions and composition of the introduction. If new populations are established from a small number of individuals and source populations, they typically have low within population genetic variation due to the effects of genetic drift during the founder event. Many rainbow trout (*Oncorhynchus mykiss*) introductions, however, have occurred from multiple sources. For instance, a study of rainbow trout from Lake Ontario utilizing microsatellite markers found highly significant levels of differentiation between populations in the lake (O'Connell *et al.* 1997). A similar allozyme study of rainbow trout populations in Lake Superior showed genetic differentiation between tributaries, although at a lesser extent than what is seen in the species' native range on the West Coast of North America (Krueger and May 1987). Previous

allozyme based studies of New Zealand trout showed no interpopulation genetic variation throughout the Central North Island; contrasting with known variation in physiology and morphology between populations (Snowdon and Adam 1992) and known physical barriers to migration such as dams and waterfalls. In fact, no published population genetic studies utilizing hyper-variable microsatellite markers have been performed on New Zealand rainbow trout. Relative to allozyme markers the more variable microsatellite DNA loci can permit greater resolution of population genetic variation (Sanchez *et al.* 1996)

Rainbow trout were first introduced from California to Lake Taupo in the late 1800's, and subsequent introductions occurred well into the early 1900's (Scott *et al.* 1978). Rainbow trout introductions were carried out for recreational and commercial purposes, and established an economically important fly fishing industry (Scrimgeour and Oxley 2001). Formed by the calderas of two volcanoes, Lake Taupo is the largest lake in New Zealand having an average depth of 100 meters and a surface area of 616 km² (Figure 3.1). Huka Falls at the lake outlet forms an impassable barrier to fish migration. The Tongariro River is the major tributary, although there are at least eleven rivers and streams which are part of the Lake Taupo catchment. Hinemaiaia River is a high-flow tributary with controlled flow due to two hydroelectric dams. The Waimarino River, however, is a natural flow river, with seasonal variation in flow and overall lower stream temperatures. The observed morphological and physiological differences between trout from the tributaries of Lake Taupo may be due to the different hydrological conditions experienced by populations living in these rivers. In these heterogeneous environments one might expect that introduced trout would have adapted to their particular spawning stream, resulting in genetic differentiation between streams, especially given salmonids potentially rapid rate of local adaptation (Fraser *et al.* 2011). In their native range there are numerous examples of salmonids harbouring local adaptations on small geographic scales (Taylor 1991). Even when introduced to new environments, salmonids can show evidence of local adaptation in as few as six generations (Fraser *et al.* 2011). Chinook salmon (*Oncorhynchus tshawytscha*) showed significant genetic differentiation from their source populations within 30 generations of introduction (Kinnison *et al.* 2002).

It is important to understand whether the spawning populations of different tributaries are genetically differentiated from each other, both to understand the population genetics of

introduced species and, more practically, to be able to manage fisheries. In this study I used microsatellite markers to determine the levels of genetic variation and structure within rivers and among rivers in the Lake Taupo catchment. The null hypothesis was that no statistical differences will be observed in population genetic variation based on geography among tributaries (i.e. panmixia throughout the catchment).

3.2 Materials and Methods

Sample collection and genetic methods

Rainbow trout were captured at ten locations throughout the Lake Taupo catchment in New Zealand by the Department of Conservation from August 2006 to March 2009 at stream traps, by angling, or with electrofishing (Figure 3.1; Table 3.2). Individuals were collected in Rotorua, approximately 80 km northeast of Lake Taupo, at Lake Tarawera in 2009 by Department of Fish and Game staff at a stream trap. Fin clips were taken and stored in 95% ethanol at 4°C until DNA extraction was performed.

Genomic DNA was extracted from fin clips using proteinase K digestion, proteins were then salted out, and the remaining pellet was subjected to two ethanol washes and resuspended in TE buffer. Some samples were also extracted using either a phenol-chloroform method (Current Protocols in Molecular Biology, 1996), or ZyGEM kits (ZyGEM Corporation) following the manufacturer's directions. DNA concentration was quantified using a spectrophotometer which examined absorbance at 280 nm (NanoDrop™ ND-1000, Thermo Scientific).

All individuals were genotyped at the following 19 microsatellite loci: *Oki23* (McClelland and Naish 2007), *Oke4* (Buchholz *et al.* 2001), *Oneu18*, *One14* (Scribner *et al.* 1996), *One102* (Olsen *et al.* 2000), *Ssa289* (McConnell *et al.* 1995), *Ssa407*, *Ssa408* (Cairney *et al.* 2000), *Ssa20.19* (Sanchez *et al.* 1996), *Ots4* (Banks *et al.* 1999), *Ots100* (Nelson and Beacham 1999), *Omy7* (K. Gharbi and R. Guyomard, INRA), *Ogo4* (Olsen *et al.* 1998), *OMM1046*, *OMM1036* (Rexroad *et al.* 2002), *Omy1011*, *Omy1001* (Spies *et al.* 2005), *Ots3M* (Greig and Banks 1999), *OmyFGT12-TUF* (Sakamoto 1996). Polymerase chain reaction (PCR) assays to amplify microsatellite alleles were performed using the protocols

developed by the SPAN loci standardization group (Stephenson *et al.* 2009) and Dr. Megan McPhee (University of Alaska Fairbanks), using four different multiplexed PCRs (Table 3.1). All DNA was amplified in a 5 µl reaction containing 2.5 µl of QIAGEN PCR master mix (QIAGEN Inc.), the specified concentration of primers per multiplex (Table 3.1), 1.5 µl of genomic DNA, and double distilled water where needed. The PCR profile consisted of an initial denaturation time of 15 minutes at 95°C followed by 28 cycles of 30 seconds at 94°C, 1 minute 30 seconds at multiplex-specific annealing temperature (Table 3.1), and 1 minute at 72°C. This was followed by an extension at 60°C. PCR products were electrophoresed on an ABI 3730 and alleles were scored using GeneMapper (ABI). Raw genotype data was binned into size ranges using the program *tandem* (Matschiner and Salzburger 2009). Three loci, *Oki23*, *Ots4* and *Ogo4*, were excluded from further analysis due to a high percentage of missing data. Therefore 16 loci were used for subsequent analysis.

Table 3.1. Microsatellite loci, fluorescent dye label used for visualization, PCR multiplex, annealing temperature and PCR concentration in a 5 µl reaction for 19 markers used in this study.

| locus | dye label | multiplex | annealing temp (°C) | PCR concentration (mM) |
|----------------------|-----------|-----------|---------------------|------------------------|
| <i>Oki23</i> | FAM | A | 51 | 0.03 |
| <i>Oke4</i> | FAM | A | 51 | 0.20 |
| <i>One14</i> | VIC | A | 51 | 0.30 |
| <i>Ssa289</i> | NED | A | 51 | 0.10 |
| <i>Ssa408</i> | PET | A | 51 | 0.20 |
| <i>Ots4</i> | FAM | B | 57 | 0.40 |
| <i>Omy7</i> | FAM | B | 57 | 0.20 |
| <i>Ogo4</i> | VIC | B | 57 | 0.20 |
| <i>One102</i> | VIC | B | 57 | 0.15 |
| <i>OMM1046</i> | NED | B | 57 | 0.15 |
| <i>Ssa407</i> | PET | B | 57 | 0.40 |
| <i>Ots100</i> | FAM | C | 57 | 0.30 |
| <i>Omy1011</i> | VIC | C | 57 | 0.16 |
| <i>Omy1001</i> | NED | C | 57 | 0.16 |
| <i>Ots3m</i> | NED | C | 57 | 0.20 |
| <i>OMM1036</i> | PET | C | 57 | 0.16 |
| <i>OmyFGT12TUF-F</i> | PET | D | 57 | 0.20 |
| <i>Oneu18</i> | NED | D | 57 | 0.20 |
| <i>Ssa20.19</i> | FAM | D | 57 | 0.20 |

Genetic variation

Conformance to Hardy-Weinberg equilibrium (HWE) and the presence of gametic disequilibrium were evaluated using GENEPOP on the web (Raymond and Rousset 1995). Estimation of exact P-values for a HWE probability test were calculated using the Markov chain method, with 1000 dememorizations, 100 batches, and 1000 iterations per batch. Locus by locus pair-wise tests for gametic disequilibrium also utilized 1000 dememorizations, 100 batches, and 1000 iterations per batch. Comparisons were considered significant if $P < 0.05$ with a sequential Bonferroni correction (Holm 1979). The presence of null alleles and large allele drop-out was assessed using the program MICROCHECKER (Van Oosterhout *et al.* 2004). Missing data was omitted from the analysis, and a 95% confidence interval was used with 1000 replications.

To determine whether sampling sufficiently represented allelic variation, allelic discovery curves were computed using a custom function in R (R development Core Team 2005). One thousand replicates of the data set were jackknifed, using an interval of one. If the allele discovery curve approaches an asymptote, it indicates that the number of samples were sufficient to represent the allelic diversity of the populations. For this analysis all populations' data were combined.

Pairwise F_{IS} , F_{ST} , and expected heterozygosity were calculated using FSTAT 2.9.3 (Goudet 1995). Allelic richness was calculated using HP-RARE v1.0 (Kalinowski 2005), which uses a rarefaction method to correct for differing sample sizes, with the smallest sample size used for calculation being $N = 12$. An exact binomial test was used to compare differences in expected heterozygosity and allelic richness. A principal components analysis was performed in GenAlEx (Peakall and Smouse 2006) in order to examine the variation between populations.

Spatial genetic structure

Geographic distance between sampling sites was estimated based on the distance by water (also known as streamwise distance). Geographic and genetic distances were compared

using a mantel test implemented in the sub-routine ISOLDE of GENEPOP on the web (Raymond and Rousset 1995), which tests for the statistical relation of two distance matrices.

An F_{ST} outlier analysis of the microsatellite loci in all populations was done using LOSITAN (Antao *et al.* 2008; Foll and Gaggiotti 2008). Average neutral F_{ST} was approximated by removing loci potentially under selection and the “force mean F_{ST} ” option was used to increase the reliability of mean F_{ST} by running a bisection algorithm over repeated simulations. Ten thousand simulations were run using an infinite alleles model and a sample size of 50. Significant outliers were determined using a 0.95 confidence interval.

Population genetic structure between populations was evaluated using the program STRUCTURE, in which K subpopulations are characterized by allele frequencies at each locus being in HWE and not in gametic disequilibrium. (Pritchard *et al.* 2000). Ten iterations were performed for each value of K, for values of K from 1 to 15. For each Monte Carlo Markov Chain iteration, a burn in of 10,000 replicates and run of 100,000 replicates of the MCMC chain were performed. The optimal value of K was assessed both by the Evanno method (Evanno *et al.* 2005) as well as by assessing the largest log likelihood (App. Figure 3.2).

Effective population size for Taupo with all populations pooled was calculated using LDNe (Waples and Do 2008), which estimates effective population size based on linkage disequilibrium data, and estimates are subsequently jackknifed to obtain confidence intervals. The program BOTTLENECK was used to test for recent population bottlenecks (Piry *et al.* 1999); departure from mutation drift equilibrium is tested by assessing heterozygosity excess or deficiency, assuming data were collected from neutral loci and that populations are experiencing mutation-drift equilibrium. This program was also used to test for a mode shift (Luikart *et al.* 1998a), an absence of low frequency alleles in a population which can indicate recent genetic bottlenecks. Variance for the two phase mutation model (TPM) was set at 30, with the proportion of the step-wise mutation model (SMM) in TPM set at 70%. 1000 iterations were performed for each population.

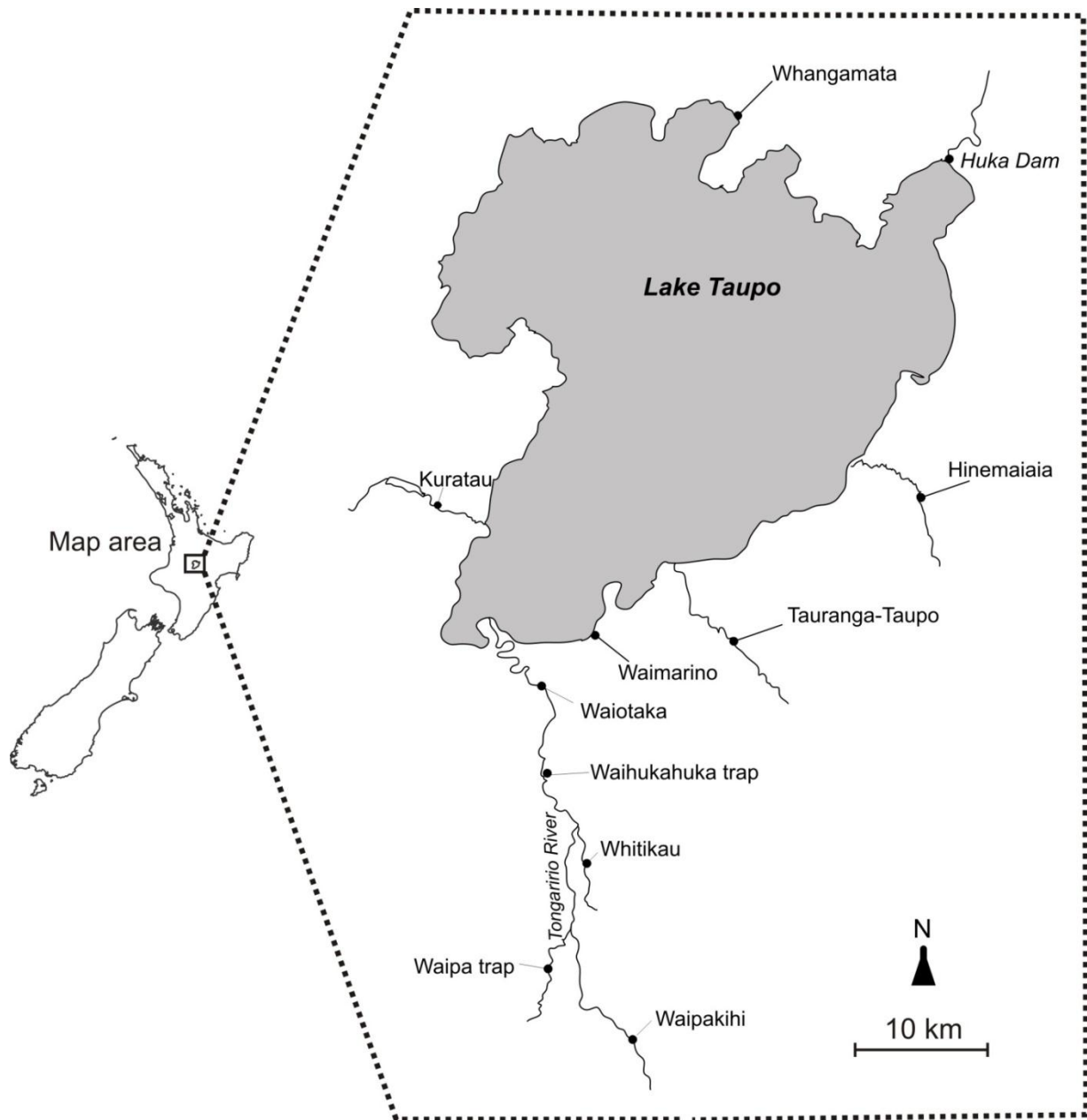


Figure 3.1. Map of ten sampling locations for rainbow trout around Lake Taupo, New Zealand. Sampling sites are named according to the tributary they are located on. Rotorua samples were collected approximately 80 km northeast of Huka Dam at the Tarawera trap.

3.3 Results

Genetic variation

All 19 microsatellite loci were polymorphic. Our allelic richness curves indicate that sampling was significant to encompass the allelic variation in the lake overall (App. Figure 3.1). MICROCHECKER detected some null alleles and allelic drop out; null alleles may be present at *Omy1001*, *Omy1011* in Hinemaiaia, *Ssa408* and *Ssa289* in Rotorua, *Ssa408* in Tauranga-Taupo, *Ssa408*, *Ots100* and *One102* in Waiotaka, *Ssa289* in Waimarino, *One102* in Kuratau, *One102*, *OmyFGT12*, and *Ssa289* in Waipa, *Ssa408*, *Ssa289*, *Ssa407* and *Ots100* in Waipakihi, and *Oneu18* in Whangamata due to homozygote excess. Stutter error may be present at *Ssa289* in Rotorua, although this is likely an artefact of the analysis since no other population showed the same result. Only the locus *One14* showed a consistent pattern of departure from HWE. After a strict Bonferroni correction no gametic disequilibrium was observed. Because of extremely high F_{IS} values (0.407 – 0.773, App. Table 3.1) and departure from HWE at 8 loci ($P < 0.001$), *One14* was excluded from further analyses. A total of 15 loci were used in the final analysis. *Ots100* and *OmyFGT12* were identified as candidates for positive selection, and *Omm1046* was detected as a candidate for balancing selection as the result of an F_{ST} outlier analysis.

Expected heterozygosity ranged from 0.722 to 0.764 in Lake Taupo, and these values were similar to H_E in the Rotorua population (0.723, Table 3.2 and App. Table 3.1). Waimarino and Waiotaka had 10 and 11 private alleles respectively, while Whangamata had no private alleles; Rotorua was within the range of Taupo values with 8 private alleles. Although the Rotorua population had an allelic richness of 4.49 which was lower than any measured Taupo population (4.54 - 4.97), it was not significantly lower than the overall Taupo allelic richness of 4.90.

Spatial genetic structure

Pairwise comparisons indicated that all populations were not significantly different at the remaining panel of 15 microsatellite loci, and that F_{ST} was low when calculated over all

populations ($F_{ST} = 0.022$) (Table 3.3). Pairwise F_{ST} values range widely, though, from 0.002 (Kuratau-Hinemaiaiaa) to 0.083 (Rotorua-Waipakihi). Overall theta is 0.022.

There was a significant positive correlation between stream distance and genetic differentiation ($r = 0.573$, $P = 0.029$), however this correlation was no longer significant when Waipakihi was removed from the analysis ($r = 0.195$, $P = 0.24$) (Figure 3.3). The FCA shows that Taupo populations cluster together with the exception of Waipakihi and Rotorua (Figure 3.2). STRUCTURE analysis indicated, using the maximum log likelihood value that the number of clusters was most likely $K=10$, and utilizing the Evanno method (Evanno *et al.* 2005) that the number of optimal clusters was likely with $K= 4$ or 15, which had the highest values of delta K (App. Figure 3.2). Upon examination none of these values of K showed clear geographic population structuring (i.e. the assignment of the majority of a geographical population to one cluster). Only the Waipa population had greater than 95% proportional membership consistently assigned in all 10 iterations to one cluster when $K = 4$, all other populations were admixed between clusters.

Seven populations were identified as departing from mutation-drift equilibrium using a Wilcoxon sign test ($P < 0.001$) in the program BOTTLENECK, these were Hinemaiaiaa, Kuratau, Tauranga-Taupo, Waiotaka, Whitikau, Waipa, and Waipakihi. However no population showed evidence of a mode shift when tested within the same program, with all populations exhibiting a classic, L-shaped distribution.

Effective population size for all of Lake Taupo was calculated as 1309.5 with a 95% confidence limit between 850.4 and 2666.8, when 0.01 is the lowest allele frequency used with a random mating model. This was the most reliable estimation, since other values of the lowest allele frequencies produced confidence intervals larger than the effective population size estimates. The variation in estimates is likely an effect of small sample size (Tallmon *et al.* 2004).

Table 3.2. Sample size, expected heterozygosity (H_E), allelic richness (Ar), private alleles (PA), and inbreeding coefficient (F_{IS}) in ten Lake Taupo and one Rotorua rainbow trout populations. Allelic richness was calculated using rarefaction to account for the smallest sample size for a locus ($N = 12$). TT represents Tauranga Taupo.

| Population | N | Ar | H_E | PA | F_{IS} |
|-------------|-----|------|-------|------|----------|
| Hinemaiaia | 40 | 4.80 | 0.748 | 6 | 0.046 |
| Kuratua | 46 | 4.70 | 0.747 | 3 | 0.010 |
| TT | 48 | 4.56 | 0.758 | 2 | 0.015 |
| Waihukahuka | 47 | 4.93 | 0.764 | 3 | 0.002 |
| Waimarino | 33 | 4.83 | 0.763 | 10 | 0.015 |
| Waiotaka | 30 | 4.97 | 0.750 | 11 | 0.156 |
| Waipa | 50 | 4.78 | 0.750 | 3 | 0.032 |
| Waipakihi | 20 | 4.63 | 0.722 | 7 | 0.114 |
| Whangamata | 25 | 4.54 | 0.722 | 0 | 0.069 |
| Whitikau | 39 | 4.96 | 0.766 | 7 | -0.019 |
| Total Taupo | 378 | 4.90 | 0.749 | 52 | 0.044 |
| Rotorua | 43 | 4.49 | 0.723 | 8 | 0.035 |

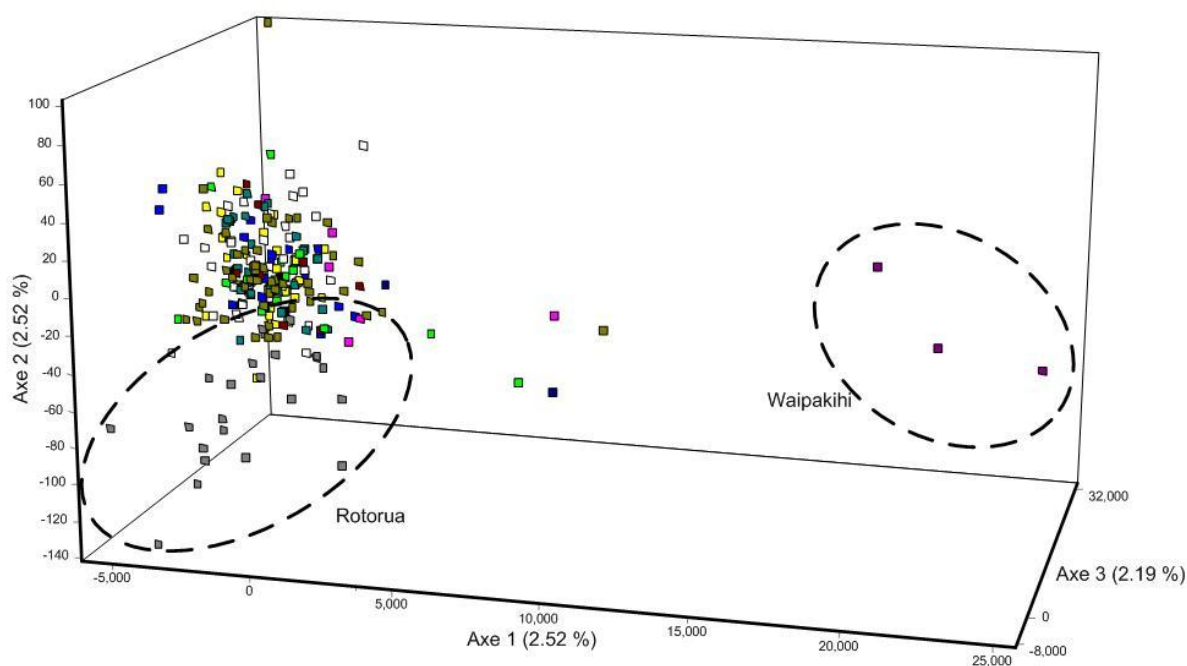


Figure 3.2. A factorial correspondence analysis of the genetic differentiation of ten riverine populations from Lake Taupo and one population from Rotorua based on allele frequencies from 14 microsatellites.

Table 3.3. Pairwise F_{ST} comparisons for 10 Lake Taupo rainbow trout populations calculated from data from 14 microsatellites. TT represents the Tauranga Taupo River. Overall theta is 0.022. P-values obtained after 1100 permutations, the adjusted significance level (5%) after a Bonferroni correction for multiple comparisons is $P < 0.001$. P values are displayed in the upper right hand corner, F_{ST} comparisons are displayed in the lower left hand corner.

| | Hinemaiaia | Kuratua | TT | Waihukahuka | Waimarino | Waiotaka | Waipa | Waipakihi | Whangamata | Whitika | Rotorua |
|-------------|------------|---------|-------|-------------|-----------|----------|-------|-----------|------------|---------|---------|
| Hinemaiaia | -- | 0.063 | 0.004 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | 0.001 | 0.002 | 0.001 |
| Kuratua | 0.002 | -- | 0.001 | 0.001 | 0.001 | 0.005 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| TT | 0.005 | 0.009 | -- | 0.002 | 0.004 | 0.010 | 0.005 | 0.001 | 0.001 | 0.002 | 0.001 |
| Waihukahuka | 0.009 | 0.009 | 0.004 | -- | 0.001 | 0.026 | 0.001 | 0.003 | 0.001 | 0.366 | 0.001 |
| Waimarino | 0.010 | 0.007 | 0.011 | 0.010 | -- | 0.038 | 0.001 | 0.002 | 0.001 | 0.001 | 0.001 |
| Waiotaka | 0.026 | 0.022 | 0.018 | 0.011 | 0.009 | -- | 0.003 | 0.096 | 0.043 | 0.014 | 0.001 |
| Waipa | 0.011 | 0.013 | 0.006 | 0.010 | 0.008 | 0.024 | -- | 0.001 | 0.001 | 0.014 | 0.001 |
| Waipakihi | 0.049 | 0.055 | 0.050 | 0.043 | 0.044 | 0.038 | 0.055 | -- | 0.015 | 0.005 | 0.003 |
| Whangamata | 0.018 | 0.020 | 0.028 | 0.030 | 0.022 | 0.042 | 0.039 | 0.063 | -- | 0.001 | 0.001 |
| Whitika | 0.007 | 0.008 | 0.006 | 0.000 | 0.009 | 0.017 | 0.006 | 0.051 | 0.027 | -- | 0.001 |
| Rotorua | 0.044 | 0.042 | 0.040 | 0.042 | 0.028 | 0.039 | 0.045 | 0.083 | 0.061 | 0.046 | -- |

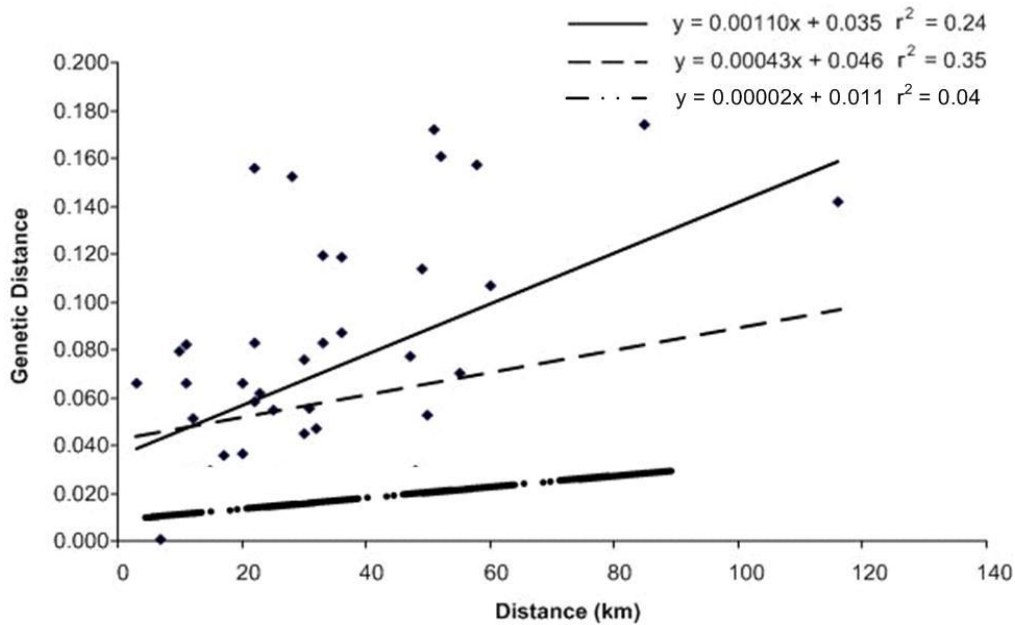


Figure 3.3. Comparison of geographic distance and Nei's genetic distance in Lake Taupo (solid line; $P = 0.029$) and within watersheds in the native steelhead range in British Columbia (dashed line; from Heath *et al.* 2001) and Taupo without the Waipakihi population (dashed and dotted line; $P = 0.024$). Data points shown are from Lake Taupo only.

3.4 Discussion

Genetic variation

Throughout this study there was no observed pattern of significant genetic differentiation among Lake Taupo populations. While overall F_{ST} values ranged widely (0.002 to 0.083), populations generally did not have a large number of private alleles, nor did they show statistically significant differences in allele frequencies. Pairwise F_{ST} comparisons (Table 3.3) revealed Taupo populations to be genetically similar, although Waipakihi and the Rotorua population did have relatively higher genetic distances when compared to all other populations. These two populations were significantly differentiated because of their geographic isolation and lack of gene flow with Lake Taupo populations. Populations within

the Lake Taupo catchment, which are geographically connected, do not show signs of genetic differentiation.

Lake Taupo has a different level of genetic variation compared to other populations of native and introduced rainbow trout that have been studied in other countries. The expected heterozygosity measured in Taupo is lower than that measured in Lake Ontario, another introduced population (0.785 to 0.952 LO, 0.722 to 0.766 Lake Taupo) and the allelic richness is significantly lower than what has been found in the anadromous ecomorph's native range in Alaska (5.68 to 10.89 Alaska, 4.56 to 4.90 Taupo) (McPhee pers. comm). The lower levels of allelic richness could be due to population bottlenecks in Lake Taupo, either as a result of the founder event during introduction, or from subsequent fluctuations in the lake's population size, and is a finding consistent with the results reported in Chapter 2. Demographic fluctuations can increase the strength of genetic drift, which eliminates some allele lineages and makes the population history difficult or impossible to ascertain from genetic data (Charlesworth *et al.* 2003). The Lake Taupo population is known to have undergone several population size contractions since its founding, although in spite of having a population size that fluctuates by a factor of four, the lowest population level was still greater than 50,000 individuals (M. Dedual, pers. comm.).

The limited amount of population differentiation in Lake Taupo may be, in part, attributable to the movement of fish by various management agencies from the time that rainbow trout introductions began in the Lake Taupo catchment. In the 1960's, 25,000 juvenile rainbow trout were released in Lake Taupo from the Lake Tarawera site in Rotorua. In addition, ova were taken from Waiotaka River and Tokaanu River rainbow trout for the main fish hatchery in Rotorua (M. Dedual, pers. comm.). This may explain the finding that some of the Rotorua samples from Tarawera cluster with individuals from several of the Lake Taupo populations. Translocations will promote gene flow and quite possibly prevent population divergence. This anthropogenic form of dispersal could initially swamp out the development of locally adapted groups and limit the effects of natural selection; however, once the genetic influences of translocation have ceased, natural selection can have a large effect on introduced individuals (Stockwell *et al.* 2003).

The genetic divergence of the Lake Taupo and Rotorua populations could also be partly due to the influence of the hatchery programme on the genetic composition of Rotorua fish. For example, Rotorua fish might experience strong genetic drift each year as a relatively small number of broodstock are used, which is artificially driving the divergence of the Rotorua population from Lake Taupo, or the possibility of hatchery selection (intentional, unintentional or natural) acting on genes that are linked to all, or some, of the microsatellite loci. Another consideration is fisheries-induced evolution, which can alter population subdivision, cause levels of genetic variation to decline, and produce selective genetic changes. Exploitative selection has previously been implicated in life-history changes to fish populations within several generations (Fraser *et al.* 2011). It is difficult to determine whether there has been a specific selective force contributing to the recent shifts in the spawning behaviour of Lake Taupo trout. One of the microsatellite loci reported in this study (*Omy12FGT-TUF*), which was a candidate for positive selection, is known to be linked to a QTL associated with spawning time (Sakamoto *et al.* 1999). It is likely that spawning time, or some other nearby linked gene function, is under selection in Lake Taupo.

Spatial genetic structure

The strong genetic structuring of many freshwater fish populations has been attributed to several factors, including natural and artificial barriers to dispersal, habitat modification, homing behaviour and site fidelity, and in some cases fishing pressure (Duncan and Lockwood 2001). We did not detect either the discrete or isolation-by-distance types of population genetic structure within Lake Taupo. The only exceptions to this pattern were the genetically divergent Waipakihi and Rotorua populations, which are known to be physically separated (i.e. there are barriers to gene flow) from Lake Taupo, and it should also be noted that Rotorua has the additional influence of a hatchery management practise. Our findings were consistent with the result of the previous study conducted by Vonlanthen and Dedual (unpublished DoC report), which found that a population sampled at Te Whai in Lake Otamangakau was significantly different from Lake Taupo. Lake Otamangakau is also physically separated from Lake Taupo. Our study showed that based on the sites sampled, Lake Taupo currently fits the model of no distinct population structure and that there is gene flow among sites (Laikre *et al.* 2005; Waples and Gaggiotti 2006). It is difficult to determine the specific amount of migration among sites, but if gene flow were low then it would not

preclude the possibility of local adaptations arising with a tributary system. The presence of private alleles in many populations, although they are at low frequency, indicated that there were some localised differences, albeit on a very small scale. Allelic measures of genetic structure are often more sensitive than measures of heterozygosity or F -statistics and this might be an early indicator of site specific changes in genetic variation (Allendorf 1986).

There are other confounding factors for determining the demographic structure of Lake Taupo rainbow trout. The size of the lake is considerably smaller than lakes studied in other countries, which have shown higher levels of genetic differentiation. For instance, Lake Ontario is 19,529 km² as compared to Taupo's 616 km², and Lake Ontario populations are genetically differentiated from each other (O'Connell *et al.* 1997). The larger habitat size could also explain the difference in estimates of effective population size, because Taupo trout have a much lower estimated N_e than Lake Ontario trout (Taupo $N_e = 1310$, Ontario $N_e = 20,012$). The estimation of N_e assumes discrete generations and panmixia (Waples 1989), and the low sample size for some populations could well have biased the estimates of N_e .

Defining a population or stock for management purposes is an important exercise for the proper management of a fishery, and molecular genetic tools offer an important resource for managers (Carvalho and Hauser 1994). While it can be conceptually complex to define a 'management unit' (Waples and Gaggiotti 2006), the traditional stock concept has been refined in recent years to include population genetic structure. This more accurately reflects fish breeding units within a harvested area and more closely aligns the management unit to a self-sustaining and demographically independent biological group (Laikre *et al.* 2005). The loss of genetic diversity as a result of a reduced population size due to either natural or human-induced changes, could have a detrimental effect on a fishery population and make it more susceptible to harvest-induced evolutionary change (Jorgensen *et al.* 2007). Maintaining levels of genetic diversity and structure through good management practises is important for the long-term sustainability of a fishery. The general rule is that managing a panmictic stock as separate stocks is generally harmless, but it is risky to manage distinct genetic stocks as a single stock. It is important to continue to monitor levels of genetic diversity and differentiation within the Lake Taupo rainbow trout population.

CHAPTER 4

Lack of spatial variation but significant seasonal variation in *OtsClock1b* in Lake Taupo rainbow trout (*Oncorhynchus mykiss*)

4.1 Introduction

Life history patterns are highly diverse in salmonids (Quinn 2005). The anadromous life history, multiple spawning runs and natal homing means that adaptation can occur at multiple life stages, including between “seasonal races”, or individuals which spawn in the same geographic location but at different times (Taylor 1991). This complexity of life histories likely evolved during periods of habitat fragmentation and divergence in isolation followed by expansion and hybridization due to the expansion and retreat of glaciers and sea-level fluctuations on the Pacific coast of western North America (reviewed in Waples *et al.* 2008). Aside from Chinook salmon (*Oncorhynchus tshawytscha*), ecotypes of *Oncorhynchus mykiss* (rainbow trout and steelhead salmon) show the most diverse array of life history strategies of all North American Pacific coast salmonids (Waples *et al.* 2001; Quinn 2005).

The key question is: how much of the observed variance in life history behaviour is attributable to genetic variation, and how much is phenotypic plasticity, which is defined as the expression of different phenotypes from the same genotype under different environmental conditions (Waples *et al.* 2004). The heritabilities of life-history traits among salmonids are often trait specific, and those heritabilities themselves may also be influenced by the quality of the environment (Carlson and Seamons 2008). Life history traits have generally lower heritability than morphological traits due to the increased selection and therefore lower genetic diversity associated with these traits (Mousseau and Roff 1987; Carlson and Williams 1999). Despite these limitations, spawning time has been shown to have a strong genetic component (Gall and Huang 1988; Siitonen and Gall 1989; Danzmann *et al.* 1994).

Although spawning time is a polygenic trait which is still not widely understood, several copies of the *Clock* gene have recently been isolated and mapped in multiple salmonids (Paibomesai *et al.* 2010). There are multiple copies of the *Clock* gene in many teleost fish due to genome duplication events, with different duplicates retained in different species (Wang 2008). In salmonids two copies were retained of *Clock1*, *Clock1a* and *Clock1b*, and one copy of *Clock3* was retained (Paibomesai *et al.* 2010). Initially the *Clock* gene was mapped in rainbow trout to a QTL that is reported to have a large effect on spawning time, development, and maturation: segregation at this QTL accounted for 50% of the phenotypic variation in males and 20% in females (Leder *et al.* 2006). This copy of the *Clock* gene was subsequently identified as *OtsClock1a* (Paibomesai *et al.* 2010). Variation in the poly-Q region of this gene is of particular interest because it has been correlated with spawning time and fitness traits in blue tit (*Cyanistes caeruleus*) (Johnsen *et al.* 2007), and acute modulation of circadian rhythms in mice (*Mus musculus*) (Vitaterna *et al.* 2006). No variation the *OtsClock1a* glutamine repeat (poly-Q region) has been found to date in Chinook salmon (*Oncorhynchus tshawytscha*) (O'Malley and Banks 2008b) or rainbow trout (Chapter 2); however an insertion-deletion polymorphism was found in this gene during its mapping (Leder *et al.* 2006), suggesting that variation in this gene might be outside the poly-Q area currently under investigation. In wild Chinook salmon populations on the west coast of North America, the pattern of *OtsClock1b* allele distribution corresponded to a latitudinal gradient and significantly differed from neutral expectations, suggesting that these differences may be maintained by selection (O'Malley and Banks 2008a). This inference of selection makes sense in light of the biological importance of spawning time for the success of offspring (Quinn *et al.* 2002).

There are several anthropogenic factors which can influence spawning time. Rearing in hatcheries often leads to earlier spawning times (Quinn *et al.* 2002), since in the wild natural selection favours later spawning to avoid redd superimposition (the nesting of one salmon on top of another salmon's nest) and disturbance (McPhee and Quinn 1998). In hatcheries early emerging fry are also protected from predation, also known as relaxed early culture selection, whereas they may be more susceptible in the wild (Quinn *et al.* 2002). Exploitation can also have an impact on spawning time, particularly since fishing pressure is often focused to only a few months of the year, often early in the year (Consuegra *et al.* 2005;

Quinn *et al.* 2006). Fishing pressure has potentially caused decreased body weight, delays in run timing, increases in age at maturation, and decreases in longevity within exploited salmon populations (Hard *et al.* 2008). It has been difficult, however, to detect genetic changes associated with these changes in phenotype (Conover and Munch 2002).

Recently Lake Taupo rainbow trout have been reported to show reduced fork length and are spawning in more concentrated runs late in the year as opposed to year-round (Figure 4.1, G. Maclean and M. Dedual pers. comm.). An autumn spawning run had been present in many Tongariro populations since introduction to the lake in the early 20th century (M. Dedual pers. comm.). The reduction of the autumn spawning run presents both a management challenge, since it is affecting angler success and hence tourism at the lake, and a scientific challenge, since the cause of the loss is unknown. It is possible that the observed shift in spawning times might be due to a selection or drift event within the Taupo catchment which has altered allele frequencies. Managers at Lake Taupo have been concerned that angling method may be selecting for specific spawning times.

This study has three aims. The first aim is to determine whether there is a spatial cline in allele frequencies in Lake Taupo populations, similar to what has been observed in west coast North American Chinook salmon (O'Malley and Banks 2008a). The second aim is to test for seasonal patterns in *OtsClock1b* allele frequencies which may be correlated with spawning run. The third aim is to investigate whether there is a genetic difference between the lake-caught and river-caught fish.

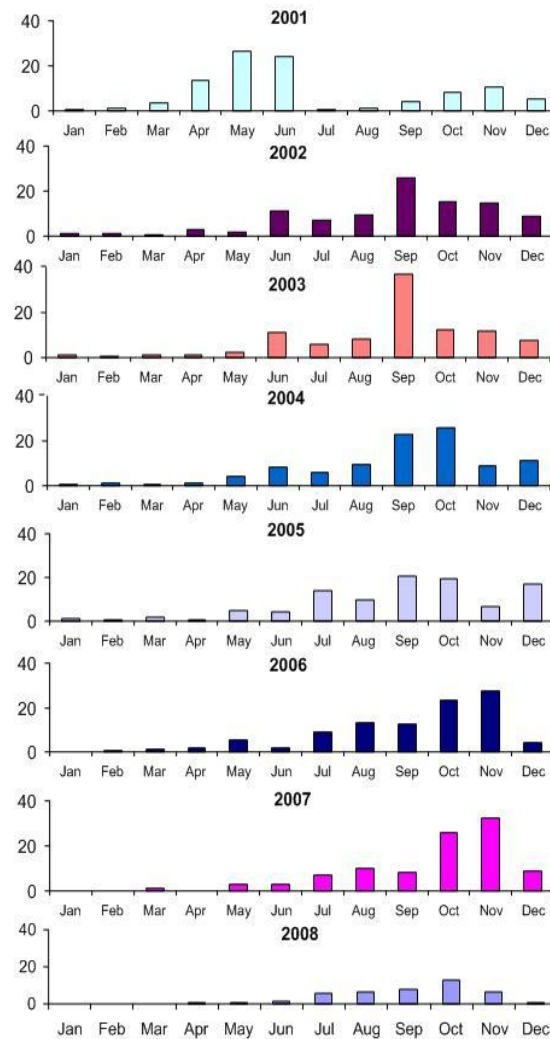


Figure 4.1. Historical changes in rainbow trout spawning time in the Waipa stream from 2001 to 2008, with number of rainbow trout trapped on the y-axis. Data from the Department of Conservation.

4.2 Materials and Methods

Spatial Sampling

Samples were collected at five representative sites around Lake Taupo to investigate the spatial pattern of variation in the timing gene *OtsClock1b* (Table 4.1). The location of these sites is detailed in Chapter 3.

Samples from each of nine California populations were analysed and compared to a selection of 50 fish from Lake Taupo (Table 4.1) to determine the genetic differences at *OtsClock1b* between Lake Taupo trout and trout from their native range in California. Samples were collected and analysed by the Southwest Fisheries Science Center.

Fishing location sampling

Fin-clips were collected from rainbow trout caught by longlining on Lake Taupo and fly-fishing on rivers to compare genetic differences at *OtsClock1b* between the two collection types. Thirty-nine river-caught trout were compared to thirty-four lake-caught fish. Fishing method was assumed to have a negligible effect (pers. comm. M. Dedual).

Seasonal sampling

Samples were collected at three sites: Waipa River trap, Waihukahuka Stream trap, and Waiotaka River at 39°0' S/175°49' E. These rivers were chosen according to the seasonal presence of fish (i.e. trout running upstream at multiple times of year), and further collection on other tributaries was not undertaken due to low numbers of fish present. Approximately 30 to 50 individuals were sampled per month, every three months at the Waipa trap and Waihukahuka trap, for autumn (May/June), winter (July, August, September), and spring (October, November, December) from 2006 to 2010. A winter (late June/July) and spring (October/November) sample were collected by angling in 2008 through 2010 for the Waiotaka (Table 4.3). These fish were measured for forklength and sexed when a fin clip was collected.

Genomic DNA extraction and genotyping

Genomic DNA was extracted using phenol-chloroform, Prep-GEM protocol (ZyGEM Corp. Limited), or Invitrogen PureLink Genomic DNA extraction kits (Life Technologies). We amplified the *OtsClock1b* gene using published primer sequences and conditions (O'Malley *et al.* 2007). The fluorescently labelled sense primer was 5'-CCTGTGTTT-GTCTCCAACAGCA-3' and the antisense primer was 5'-CTGTCACTGCGAAATTACAGTCCT-3. DNA was amplified in 10 µl reactions using 0.20 mM of the sense and antisense primer, 2.5 mM MgCl₂, 1 X buffer, 0.20 mM dNTPs, and 0.5 units of Taq. The touchdown

PCR profile consisted of one initial denaturing cycle of 3 min at 94°, followed by one cycle of 1 min at 94 °C, 1 min at 62 °C annealing temperature and 1 min 30 s at 72 °C. In subsequent cycles, the annealing temperature was decreased by 2° until 56 °C was reached, followed by 29 more cycles of 1 min 94 °C, 1 min at 56 °C, 1 min 30 s at 72 °C and a final extension of 10 min at 72 °C. PCR products were initially electrophoresed on an ABI 3730 and analysed as length polymorphisms using GENEMAPPER software. Subsequently PCR product was electrophoresed on 2.5% fine-sieve agarose gels at 100 volts for 45 minutes, using a 50 bp ladder for fragment length identification following staining with ethidium bromide. A total of 700 fish were analysed with a mean of 87 fish per site and season (Table 4.1).

Table 4.1. Samples used to investigate spatial variation in the *OtsClock1b* gene in both Lake Taupo and California.

| site | N | Date of collection |
|-----------------------|----|-----------------------------------|
| <i>Lake Taupo</i> | | |
| Whitikau River | 35 | October 2006 |
| Waimarino River | 34 | October 2008 |
| Kuratau River | 34 | August 2006 |
| Taruranga-Taupo River | 41 | June-July 2007 |
| Hinemaiaia River | 31 | August 2006 |
| <i>California</i> | | |
| Lake Almanor | 51 | November 2002 |
| American River | 53 | September, October, December 2005 |
| Big Creek | 15 | November 2003 |
| Eel River | 15 | February 1999 |
| Middle Fork Eel River | 37 | August 1999 |
| Lagunitas Creek | 61 | August 2001 |
| McCloud River | 53 | November 2002 |
| Miller Creek | 70 | October 2001 |
| Scott Creek | 16 | March 2000 |

Two samples from the Waipa trap from November 2006 and May/June 2007 were also genotyped at 12 neutral microsatellite loci: *Oke4* (Buchholz *et al.* 2001), *Oneu18*, *One14* (Scribner *et al.* 1996), *Ssa289* (McConnell *et al.* 1995), *Ssa408* (Cairney *et al.* 2000), *Ssa20.19* (Sanchez *et al.* 1996), *Ots100* (Nelson and Beacham 1999), *OMM1036* (Rexroad *et al.* 2002), *Omy1011*, *Omy1001* (Spies *et al.* 2005), *Ots3M* (Greig and Banks 1999), and

OmyFGT12-TUF (Sakamoto 1996). Polymerase chain reaction (PCR) assays to amplify microsatellite alleles were performed using the protocols developed by the SPAN loci standardization group (Stephenson *et al.* 2009) and Dr. Megan McPhee (University of Alaska Fairbanks), using four different multiplexed PCRs (Table 4.2). All DNA was amplified in a 5 µl reaction containing 2.5 µl of QIAGEN PCR master mix (QIAGEN Inc.), the specified concentration of primers per multiplex (Table 4.2), 1.5 µl of genomic DNA, and double distilled water where needed. The PCR profile consisted of an initial denaturation time of 15 minutes at 95°C followed by 28 cycles of 30 seconds at 94°C, 1 minute 30 seconds at multiplex-specific annealing temperature (Table 4.2), and 1 minute at 72°C. This was followed by an extension at 60°C. PCR products were electrophoresed on an ABI 3730 and alleles were scored using GeneMapper (ABI).

Table 4.2. Microsatellite loci, fluorescent dye label used for visualization, PCR multiplex, annealing temperature and PCR concentration in a 5 µl reaction for 12 neutral markers used in this study.

| locus | dye label | multiplex | annealing temp (°C) | PCR concentration (mM) |
|--------------------|-----------|-----------|---------------------|------------------------|
| <i>Oki23</i> | FAM | A | 51 | 0.03 |
| <i>Oke4</i> | FAM | A | 51 | 0.20 |
| <i>One14</i> | VIC | A | 51 | 0.30 |
| <i>Ssa289</i> | NED | A | 51 | 0.10 |
| <i>Ssa408</i> | PET | A | 51 | 0.20 |
| <i>Ots100</i> | FAM | C | 57 | 0.30 |
| <i>Omy1011</i> | VIC | C | 57 | 0.16 |
| <i>Ots3M</i> | NED | C | 57 | 0.20 |
| <i>OMM1036</i> | PET | C | 57 | 0.16 |
| <i>OmyFGT12TUF</i> | PET | D | 57 | 0.20 |
| <i>Oneu18</i> | NED | D | 57 | 0.20 |
| <i>Ssa20.19</i> | FAM | D | 57 | 0.20 |

Statistical analysis: Spatial variation

Exact tests for genic differentiation (G test) were conducted using GENEPOP on the web (Raymond and Rousset 1995) to assess the statistical difference of allele frequencies between spatial sites and for comparisons to the native range in California. Statistical tests were carried out to determine the genetic differentiation at *OtsClock1b*. Pairwise F_{ST} values between collections were calculated using FSTAT 2.9.3 (Goudet 1995) for spatial comparisons, comparisons to the native range, and between Taupo samples.

Statistical analysis: Temporal variation

For temporal samples, Hardy-Weinberg equilibrium was tested using GENEPOP on the web (Raymond and Rousset 1995) utilizing a probability test, a test for heterozygote excess, and a test for homozygote excess with estimation of P-Values by the Markov chain method, using 1000 dememorizations, 100 batches and 1000 iterations per batch. Exact tests for genic differentiation (G test) were also conducted using GENEPOP on the web (Raymond and Rousset 1995) to assess the statistical difference of allele frequencies between seasons at each site. Pairwise F_{ST} values between collections were calculated using FSTAT 2.9.3 (Goudet 1995).

Table 4.3. The number (N) of male and female rainbow trout genotyped from seasonal samples at three tributaries of Lake Taupo.

| | | | | | | | | | |
|--------|---------------------|--------|-------|-------------------------|--------|-------|--------------------|--------|-------|
| trap | Waihukahuka (HS) | | | | | | | | |
| season | Autumn (2009, 2010) | | | Winter (2006,2009,2010) | | | Spring (2009,2010) | | |
| sex | Male | Female | Total | Male | Female | Total | Male | Female | Total |
| N | 24 | 36 | 60 | 44 | 30 | 74 | 27 | 47 | 74 |

| | | | | | | |
|---------|--------------------|--------|-------|-------------------------|--------|-------|
| angling | Waiotaka (WA) | | | | | |
| season | Winter (2009,2010) | | | Spring (2008,2009,2010) | | |
| sex | Male | Female | Total | Male | Female | Total |
| N | 26 | 18 | 44 | 52 | 34 | 86 |

| | | | | | | | | | |
|--------|-------------------------|--------|-------|-------------------------|--------|-------|-------------------------|--------|-------|
| trap | Waipa (WP) | | | | | | | | |
| season | Autumn (2007,2009,2010) | | | Winter (2006,2008-2010) | | | Spring (2006,2008-2010) | | |
| sex | Male | Female | Total | Male | Female | Total | Male | Female | Total |
| N | 50 | 54 | 104 | 72 | 58 | 130 | 54 | 74 | 128 |

A generalized linear model (GLM) using the stepwise backward elimination method in R (see App. Figure 4.1; Ihaka and Gentleman 1996; Team 2011) was used to test the interaction of phenotypic factors (forklength and sex) and location with allele frequency. Since the data were overdispersed, a quasi-binomial distribution was used for statistical testing. A global model including multiple variables and interactions was constructed, and then interactions and variables removed from the model. If an analysis of deviance, utilizing an F -statistic test, showed that the simpler model was not significantly different from the

more complex model, then the model was further simplified, until only significant factors remained. The residual deviance was then used to perform a goodness-of-fit test.

Statistical analysis: Fishing location

Exact tests for genic differentiation (G test) were conducted using GENEPOP on the web (Raymond and Rousset 1995) to assess the statistical difference of allele frequencies for the two different fishing types.

4.3 Results

OtsClock1b allele frequency distribution geographically

Two alleles were found at the *OtsClock1b* gene, 324 and 345 base pairs respectively. Spatial samples around Lake Taupo show little divergence at *OtsClock1b*, with only Tauranga-Taupo and Whitiakau being significantly different ($\chi^2 = 7.84$, $P = 0.02$, Figure 4.2), similar to the low divergence at the microsatellite loci (Chapter 3).

Lake Taupo trout have a significantly higher frequency of the 345 bp allele than their putative source populations in California ($\chi^2 \geq 9.35$, $P < 0.01$, Figure 4.3). While no Californian population had the 345 bp allele at a frequency higher than 6% in the population, an average of several Lake Taupo populations had an allele frequency of 29.8%, and in the timing samples taken in Lake Taupo the frequency of the 345 bp allele ranges between 26% and 39% (Figure 4.6).

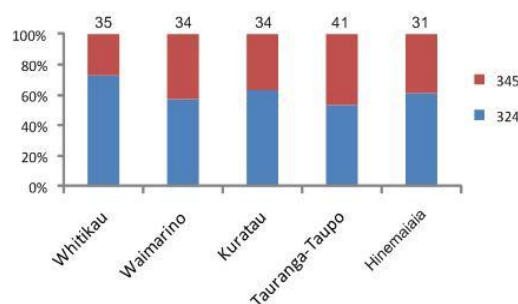


Figure 4.2. Spatial distribution of *OtsClock1b* allele frequencies at five sampling sites around Lake Taupo. Sample size is noted at the top of the bars. There is only significant genic difference between Tauranga-Taupo and Whitiakau ($\chi^2 = 7.84$, $P = 0.02$).

OtsClock1b allele frequency distribution seasonally

Seasonal variation in allele frequency at *OtsClock1b* appeared to be site specific. The Waipa trap showed a gradual increase in the frequency of the 345 bp allele from autumn through winter and into spring (Figure 4.6). A comparison of pairwise F_{ST} values indicated that the genetic difference between seasonal samples at *OtsClock1b* was greater than at neutral markers (Figure 4.4). There was a significant genic difference between Waipa autumn and the other Waipa samples ($\chi^2 > 15$, $P < 0.001$). There was a statistical trend between the Waihukahuka autumn and winter samples ($\chi^2 = 5.76$, $P = 0.055$). The Waiotaka sample also appears to show a slight increase in 345 bp allele frequencies from winter to spring. The winter Waiotaka sample is the only sample which is not in Hardy Weinberg equilibrium, because it contained no homozygotes for the 345 bp allele ($P = 0.005$). The Waihukahuka, however, showed an increase in the winter but then a slight decrease in spring. This same pattern was reflected at each site in the genotypic frequencies (Figure 4.7). Seasonal differences in genotypic frequency at *OtsClock1b* at three sampling sites around Lake Taupo. There was a significant genotypic difference between Waipa autumn and other Waipa samples ($\chi^2 > 15$, $P < 0.001$).

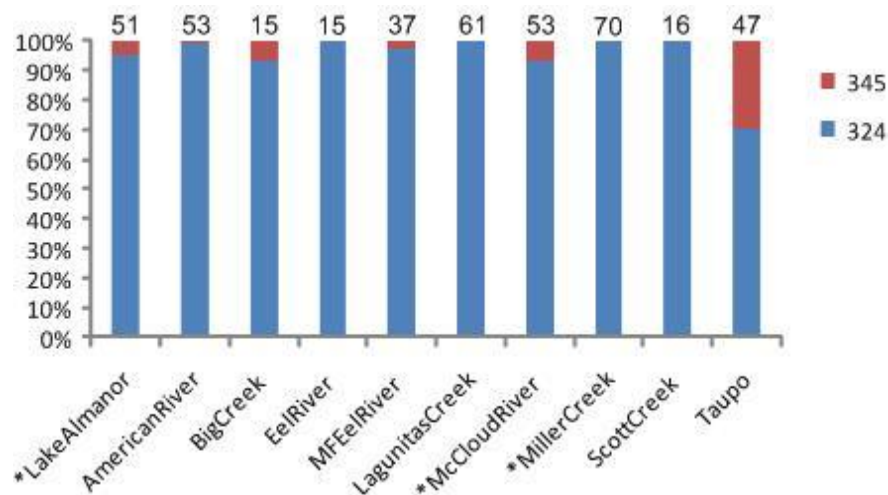


Figure 4.3. Allele frequencies at *OtsClock1b* in nine Californian populations and a combined Lake Taupo sample from several different tributaries. An asteriks denotes a putative source population for the Lake Taupo introduction. Sample size is noted at the top of the bars. Lake Taupo is significantly different from all other populations ($\chi^2 \geq 9.35$, $P < 0.01$).

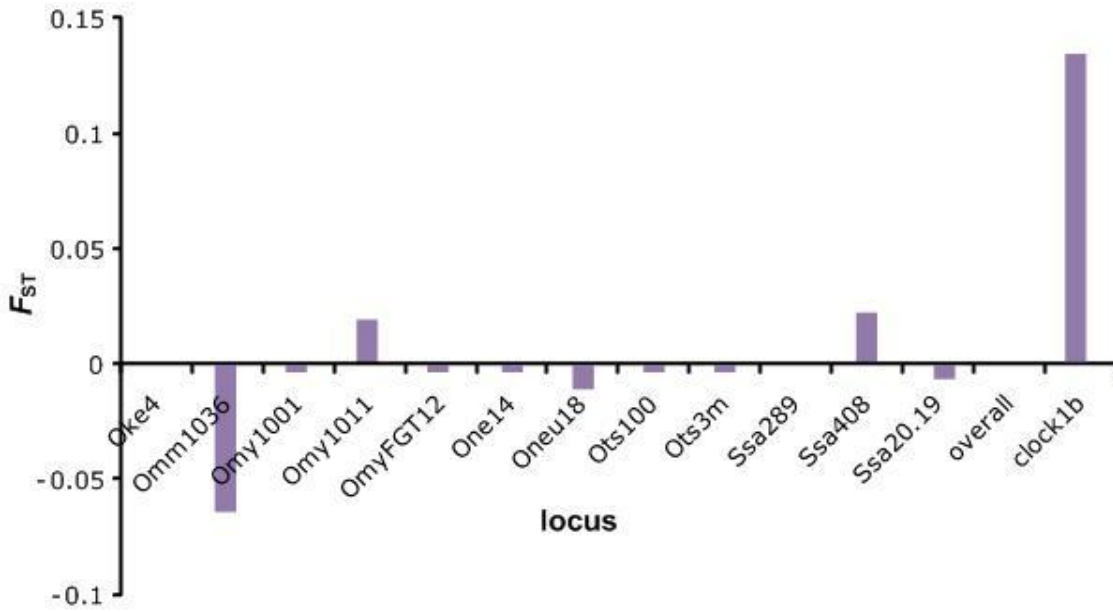


Figure 4.4. Pairwise F_{ST} values for the November 2006 and May/June 2007 genetic samples taken at Waipa trap.

Pairwise F_{ST} comparisons between seasonal samples showed significant differences between the Waipa autumn sample and Waipa winter and spring samples, and between the early season Waipa sample and later season Waihukahuka samples (Table 4.4).

Table 4.4. Pairwise F_{ST} values for seasonal samples from Lake Taupo – see Table 4.1 for site abbreviations. P -values obtained after 560 permutations, the adjusted significance level (5%) after a Bonferroni correction for multiple comparisons is $P = 0.002$. P -values are displayed in the upper right hand corner, F_{ST} comparisons are displayed in the lower left hand corner. Significant values are bolded.

| | HSaut | HSwint | HSspr | WAwint | WAspr | WPaut | WPwint | WPspr |
|--------|--------------|--------|--------------|--------|--------------|--------------|--------------|--------------|
| HSaut | -- | 0.070 | 0.171 | 0.245 | 0.011 | 0.239 | 0.025 | 0.002 |
| HSwint | 0.019 | -- | 0.613 | 0.648 | 0.570 | 0.004 | 0.886 | 0.141 |
| HSspr | 0.008 | 0.000 | -- | 1.000 | 0.232 | 0.002 | 0.504 | 0.046 |
| WAwint | 0.005 | 0.000 | 0.000 | -- | 0.270 | 0.016 | 0.493 | 0.043 |
| WAspr | 0.039 | 0.000 | 0.003 | 0.003 | -- | 0.002 | 0.604 | 0.420 |
| WPaut | 0.004 | 0.065 | 0.045 | 0.042 | 0.092 | -- | 0.002 | 0.002 |
| WPwint | 0.024 | 0.000 | 0.000 | 0.000 | 0.000 | 0.068 | -- | 0.113 |
| WPspr | 0.064 | 0.007 | 0.019 | 0.020 | 0.000 | 0.123 | 0.006 | -- |

Table 4.5. Summary of stepwise backward elimination of factors model selection for 345 bp allele frequency in Lake Taupo rainbow trout. $P < 0.01$ necessary for model inclusion. Final model and goodness-of-fit are reported.

| Model | Residual deviance | df | Δ deviance | Δ df | Probability |
|-----------------|-------------------|----|-------------------|-------------|-------------|
| Null | 4.27 | 4 | | | |
| -season:sex | 9.75 | 6 | 5.48 | 2 | 0.158 |
| -trap:sex | 16.0 | 10 | 6.27 | 4 | 0.455 |
| -trap:season | 19.9 | 12 | 3.93 | 2 | 0.321 |
| -sex | 20.8 | 13 | 0.86 | 1 | 0.475 |
| -trap | 21.7 | 15 | 0.88 | 2 | 0.756 |
| + trap, -season | 51.2 | 15 | 30.4 | 2 | 0.002 |

Final model: 345allele~season, probability = 0.116

The relative effects of several factors were tested with a stepwise backwards elimination of factors in generalized linear models. The results of these tests indicated that sex, trap and forklength were not significant factors in the difference in allele frequency, but that season was a significant factor (Table 4.5). Subsequently the residual deviance for the final model was used, with only season as a factor, to perform a goodness-of-fit test. The goodness of fit test was not significant ($P = 0.116$) indicating that the model fits the data. The estimated coefficients for the seasons were -0.6581 for autumn, 0.8999 for winter, and 1.1043 for spring respectively. Taking the exponential of estimated coefficients of the later two seasons indicates the probability of seeing the 345 bp allele in those seasons in relation to autumn, the intercept. This indicates that there is a 2.46 greater chance of seeing the 345 bp allele in winter and a 3.02 times greater chance of seeing the 345 bp allele in spring than in autumn respectively.

Effects of fishing location

There did not appear to be a difference in *OtsClock1b* genotypes between collection locations, since fish caught by trolling in the lake, and fish caught in streams by fly fishing had similar allele frequencies ($\chi^2 = 1.12$, $P = 0.568$, Figure 4.5).

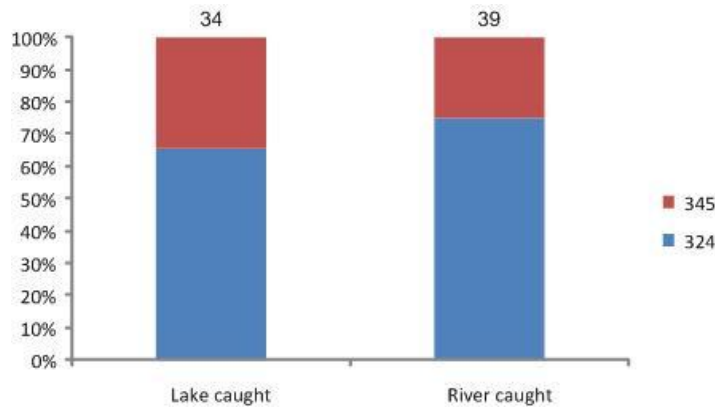


Figure 4.5. Allele frequencies at *OtsClock1b* compared between two different collection locations taken in winter of 2009. Frequencies are not significantly different ($\chi^2 = 1.12$, $P = 0.568$). Sample size is noted at the top of the bars.

4.4 Discussion

In this study I investigated *OtsClock1b* variation in Lake Taupo rainbow trout spatially, seasonally, and between different fishing methods. While there was no evidence for geographic structuring or a difference in the effect of different fishing locations, there was some evidence that allele frequencies varied seasonally.

OtsClock1b allele frequency distribution geographically

Only the 324 bp allele at *OtsClock1b* has been observed in *O. mykiss* samples from the northwest coast of the United States (K. O'Malley pers. comm), but this study found a second larger allele, at 345 bp, in both Lake Taupo populations and in Californian populations. It has been hypothesized that larger alleles in the poly-Q region of *Clock* genes have been associated with seasonal rate-of-change of photoperiod since higher latitudes experience a different rate of change than lower latitudes (Johnsen *et al.* 2007; O'Malley and Banks 2008a). Our study area is much smaller than those shown to have latitudinal gradients

in previous studies, and accordingly there was not a latitudinal pattern to the allele frequency variation discovered in the *OtsClock1b* gene.

The 345 bp allele in Lake Taupo rainbow trout was most likely introduced to New Zealand in the founder populations sourced from California. It is less likely to be a result of a parallel mutation given the limited time following the introduction. However the Lake Taupo population did show a much higher frequency of this large allele, which was either caused by a founder effect (i.e. more individuals with the 345 bp allele were translocated from California), natural selection for the larger allele over time, or the chance effect of genetic drift since introduction. Changes in the poly-Q region alter the circadian phenotype by affecting binding affinity of the *Clock* transcription factor (Darlington *et al.* 1998), and because of this direct relationship to phenotype *Clock* could be under positive selection. Therefore it is unlikely that a change in allele frequency would be caused by random genetic drift, so genetic differences at the *OtsClock1b* gene are either due to a founder effect or natural selection.

OtsClock1b allele frequency distribution seasonally

Season was the only significant factor affecting allele frequency at *OtsClock1b*, and that the frequency of the 345 bp allele was significantly less in autumn at the Waipa trap, as was the frequency of homozygotes for the 345 bp allele. While there was a decrease in the presence of the 345 bp allele at the Waihukahuka trap in autumn, the difference between seasons at this site showed a trend but was not significant ($P = 0.056$). Recently a similar pattern has also been discovered in blue tits, where female birds with fewer poly-Q repeats laid eggs earlier in the season, which suggests that the *Clock* gene might be a component of local adaptation to seasonal environments (Liedvogel *et al.* 2009).

Spawning time is heritable in rainbow trout (Siitonen and Gall 1989). It is possible that the variation in the data and the intermittent pattern of seasonal variation between sites is due to phenotypic plasticity. Modifications of spawning behaviour in response to the environment may be a way to mitigate the effects of interannual environmental variation on reproductive success (Quinn and Adams 1996); therefore some flexibility of the spawning time trait was likely selected for over years of salmonid evolution (Waples *et al.* 2008). The

study of introduced Chinook salmon in New Zealand has shown that local populations can also evolve divergent life-history types even though they come from a recent common ancestral source (Kinnison *et al.* 2011). In the absence of a “common garden” experiment it is difficult to say how much of the site variation observed in this study is due to aspects of sampling or environment, and how much of the variation is due to genetic differences between sites. In South Island New Zealand Chinook salmon, though, seasonal variation has actually led to neutral differentiation as well (Quinn *et al.* 2000). Ultimately the seasonal differentiation observed at Waipa trap could lead to neutral differentiation between different temporal “populations” based on spawning time. However, sites closer to Lake Taupo, like Waihukahuka and Waimarino may have hatchery introgression or other confounding factors that prevent this sort of differentiation.

Effect of angling: exploitative selection?

The variability of life history characteristics is an important factor in a species’ response to exploitative selection (Dunlop *et al.* 2007), but human exploitation, including fishing pressure, often changes phenotype more rapidly than other agents of change because of the intensity of selection for particular traits (Darimont *et al.* 2009). Selective exploitation by angling of early running fish has altered phenotypic and genetic traits in Iberian Atlantic salmon (Consuegra *et al.* 2005). Angling can exert sufficient exploitative pressure to alter the life-history traits of fish stocks, particularly migration (Theriault *et al.* 2008). A difference in the genotypes being caught in different fishing habitats was not detected, so if there was exploitative selection, fishing in different habitats does not seem to be necessarily selecting for a particular spawning time. Fishing effort through the spawning season (May through November) is at its most stable for the year, with wider fluctuations in summer than in winter. This is true for both the lake and river habitats, which diminishes the possibility that the observed spawn timing shift is due to intense season specific exploitation (Figure 4.8). However the combined lake and river fishing pressure may be having a large overall impact on the population. The selective exploitation of certain population components by anglers can lead to an overall decrease in population viability due to a loss of genetic variation and decrease in body size (Consuegra *et al.* 2005; Quinn *et al.* 2006). Conversely, some salmon stocks have shown resilience to exploitation (Quinn *et al.* 2007), and it is possible that the changes in the Taupo fishery are attributable to phenotypic plasticity. Genetic monitoring is a

valuable tool to understand the genetic impact of exploitation (Allendorf *et al.* 2008), and further genetic monitoring is certainly warranted to fully understand the causes of the Lake Taupo early-run decline.

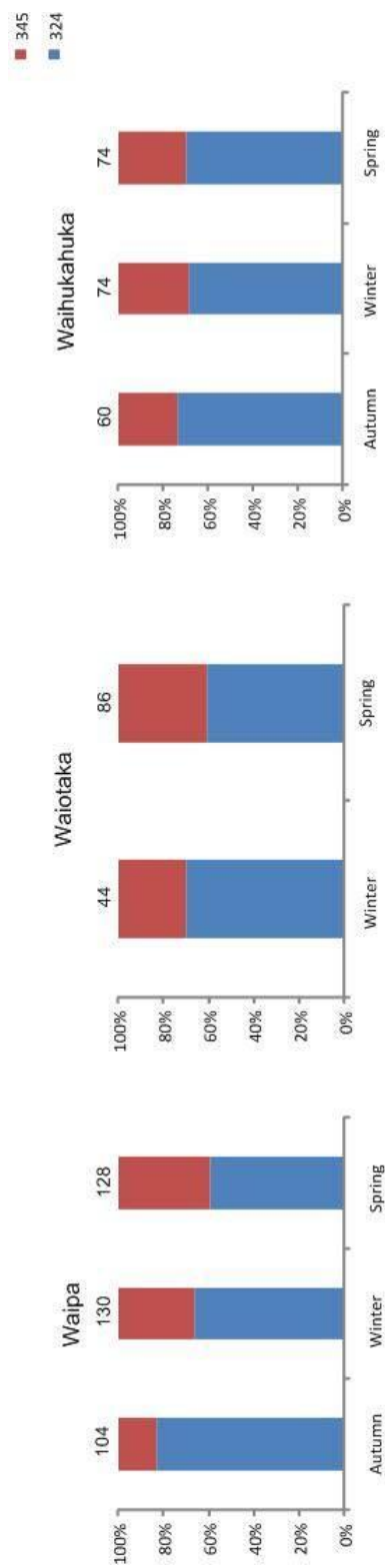


Figure 4.6. Seasonal differences in allele frequencies at *OsClock/b* at three sampling sites around Lake Taupo. Sample size is noted at the top of the bars. There is a significant genic difference between Waipa autumn and the other Waipa samples ($\chi^2 > 15$, $P < 0.001$). There is a statistical trend between the Waihukahuka autumn and winter samples ($\chi^2 = 5.76$, $P = 0.055$).

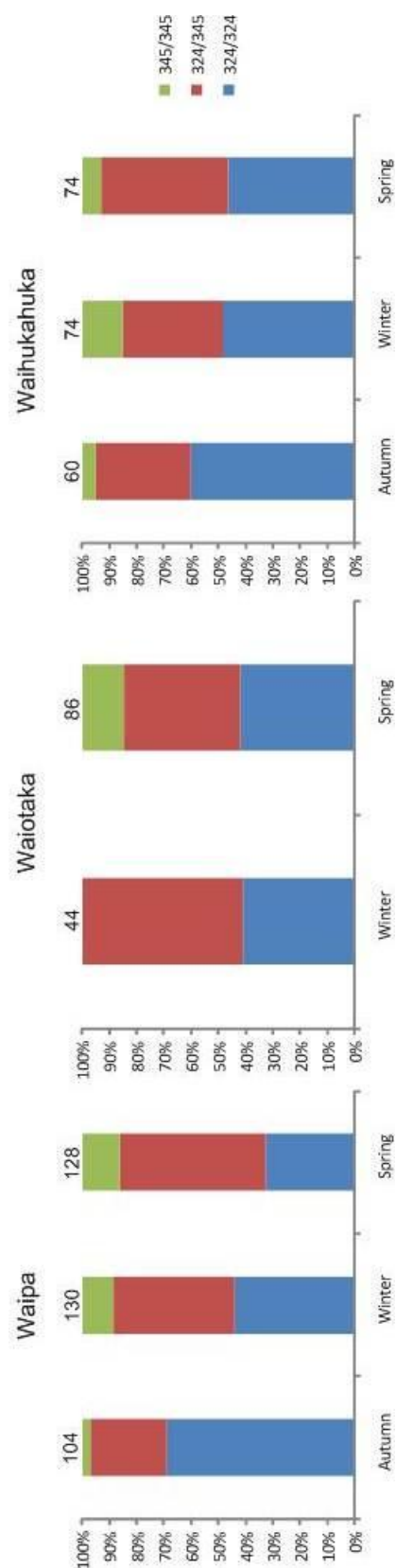


Figure 4.7. Seasonal differences in genotypic frequency at *OtsClock1b* at three sampling sites around Lake Taupo. Sample size is noted at the top of the bars. There is a significant genotypic difference between Waipa autumn and other Waipa samples ($\chi^2 > 15$, $P < 0.001$).

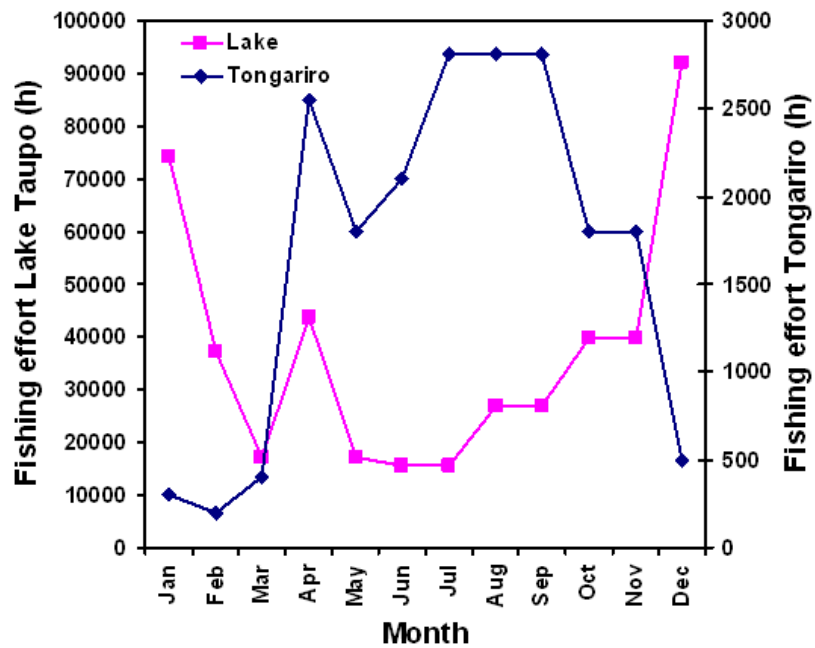


Figure 4.8. Fishing effort on Lake Taupo and the Tongariro River for 2005 in hours. Figure courtesy Department of Conservation Taupo-nui-a-Tia Office, Turangi.

CHAPTER 5

Minimal temporal genetic variation in Lake Taupo rainbow trout

Population genetics studies are frequently only a snapshot in time, analysing data from a sample of one year's cohort, and, depending on sampling strategy, this can lead to an incorrect assessment of population structure because it fails to take into account temporal variation in genetic diversity and population differentiation (Waples 1998). Without the analysis of historical samples the effect that temporal variation has on the results of a study on spatial population structure is not quantified. Studies often interpret outlier populations in population genetic structure analysis as evidence for genetic bottlenecks or outcrossing events (Heath *et al.* 2001), which assumes that genetic diversity has changed relatively recently as the result of a demographic shifts (Queney *et al.* 2000). Inferences from such studies provide greater statistical power when assumptions about the processes that have led to genetic structure can be verified using temporal samples (Nielsen *et al.* 1997; Nielsen *et al.* 1999a; Tessier and Bernatchez 1999; Queney *et al.* 2000; Heath *et al.* 2002). However the availability of historical samples to study is rare.

Genetic drift can cause larger stochastic fluctuations in allele frequencies in small populations – in salmonids, variation in the number of spawners in a year directly affect the magnitude of the fluctuations (Waples and Teel 1990). Population bottlenecks result in the loss of genetic diversity (Luikart and Cornuet 1998) and often leave a population more susceptible to the effects of genetic drift due to a reduction in effective population size. For instance, Atlantic salmon (*Salmo salar*) show a decline in the number of alleles over time, likely as a consequence of exploitation (Nielsen *et al.* 1997). The loss of alleles due to genetic drift reported in New Zealand Snapper (*Pagrus auratus*) lead to low genetic diversity in current populations after years of exploitation (Hauser *et al.* 2002). Although demographic changes significantly affect genetic diversity within populations, genetic differentiation between populations can remain relatively stable over time even after a large population

bottleneck. Stable patterns of genetic differentiation of populations over time have been shown in landlocked Atlantic salmon (Tessier and Bernatchez 1999), peregrine falcons (*Falco peregrines*) (Brown *et al.* 2007), Newfoundland cod (*Gadus morhua*) (Ruzzante *et al.* 2001), and European otters (*Lutra lutra*) (Pertoldi *et al.* 2001). In cod populations genetic differentiation is thought to be preventing recovery of populations since it signals a lack of migration (Ruzzante *et al.* 2001), which could increase effective population size and replenish genetic variation among populations. The maintenance of genetic differentiation during population bottlenecks does not always result in a loss of genetic diversity, though, as shown in peregrine falcons, which actually exhibit more contemporary genetic diversity than that found in historical samples taken before their DDT-induced bottleneck (Brown *et al.* 2007). This is due to the introduction of individuals from fostering programs, some of which came from extra-continental populations. North Sea sole are another example of the maintenance of genetic diversity in spite of a population decline (*Solea solea*), since they have consistent levels of genetic diversity despite exploitation over the last fifty years (Cuveliers *et al.* 2011). Despite a reduction in population size it seems that genetic recovery of diversity can happen on shorter time scales than previously thought. Effective population size (i.e. the number of reproductive individuals) is not the only potential cause of temporal fluctuation in genetic diversity and differentiation, though. Demographic stochasticity, particularly with regard to reproductive success between individuals, can also lead to changes in allele frequencies over time.

Several studies have examined temporal variation in genetic structure of salmonids. In wild steelhead in British Columbia, Heath *et al.* (2002) found little or no change in levels of genetic diversity but large variation in genetic structure over time, in spite of some highly structured populations. On the other hand, Hansen (2002) found that historical and contemporary samples of brown trout were genetically similar despite the stocking of domesticated trout, although admixture with indigenous fish was prevalent over time. Therefore the temporal stability of genetic structure in salmon populations is largely reliant on the environment and their history of management.

Collections of historic fish scales have been maintained at many fisheries institutions throughout Europe and North America for fish ageing purposes. There is often a small piece of desiccated tissue still attached to these scales which contains enough historic DNA for

analysis. The samples were often dried quickly and there is a limited occurrence of tissue degradation or decay (Nielsen *et al.* 1999a). Several protocols have been developed for the extraction of DNA from fish scales (Nielsen *et al.* 1999a), although scales that are less than fifty years old are often successfully processed with standard DNA extraction techniques (Eric Iwamoto pers. comm.).

This study is investigating the introduced population of rainbow trout (*Oncorhynchus mykiss*) in Lake Taupo New Zealand, which may have lower levels of diversity at neutral loci compared with native source populations (Chapter 2). It is unknown whether this lower diversity is due to genetic drift resulting from the bottleneck that occurred as a result of introduction or a subsequent reduction of population size, or if it is due to a selective sweep within Lake Taupo. The investigation of historical samples will help to elucidate whether understand recent local events in the Lake Taupo rainbow trout population.

The aim of this study was to determine whether there has been temporal variation in the population genetic structure and levels of diversity of Lake Taupo rainbow trout. I also examined a candidate gene for selection, *OtsClock1b*, to determine if there has been a change over time at this locus. The null hypothesis is that there has been no change over time, in either genetic diversity, genetic structure, or in the spawning time locus *OtsClock1b*.

5.2 Materials and Methods

Genomic DNA extraction and genotyping

Scales were collected by Marvin Rosenau from Tokaanu stream, Waimarino Stream, and Hinemaiaia Stream in September 1987 in the Lake Taupo catchment as part of his PhD research at Waikato University (Rosenau 1991). Samples were stored in a dry dark place at room temperature. Scale samples were mounted on slides with glycerine, and were subsequently removed with sterile razor blades.

All work with historic samples pre-PCR was done in facilities physically separated from where tissue and post-PCR materials were processed, and lab staff did not visit the historic lab if they had worked with tissue on the same day. Genomic DNA was extracted using Invitrogen PureLink Genomic DNA extraction kits (Life Technologies) with the tissue

protocol, with the scale removed after the digestion step. DNA concentration was quantified using a spectrophotometer which examined absorbance at 280 nm (NanoDrop™ ND-1000, Thermo Scientific). Details of the samples are summarized in Table 5.2.

Tissue was collected in 2006 and 2008 in Hinemaiaia River and Waimarino River respectively. Fin clips were preserved in 70% ethanol which was drained and refreshed within a week of collection. Genomic DNA was extracted from fin clips using proteinase K digestion, proteins were then salted out, and the remaining pellet was subjected to two ethanol washes and resuspended in TE buffer.

All individuals were genotyped at the following seven microsatellite loci: *Ots100* (Nelson and Beacham 1999), *Omm1046* (Rexroad *et al.* 2002), *Ssa407* (Cairney *et al.* 2000), *Omy1011*, *Omy1001* (Spies *et al.* 2005), *Ots3M* (Greig and Banks 1999), *OmyFGT12-TUF* (Sakamoto 1996). Polymerase chain reaction (PCR) assays to amplify microsatellite alleles were performed using the protocols developed by the SPAN loci standardization group (Stephenson *et al.* 2009) and Dr. Megan McPhee (University of Alaska Fairbanks). All PCRs were performed in a 5 µl reaction. All DNA was amplified in a 5 µl reaction containing 2.5 µl of QIAGEN PCR master mix (QIAGEN Inc.), the specified concentration of primer (Table 5.1), 1.5 µl of genomic DNA, and double-distilled water where needed. The PCR profile consisted of an initial denaturation time of 15 minutes at 95 °C followed by 28 cycles of 30 seconds at 94 °C, 1 minute 30 seconds at 57 °C, and 1 minute at 72 °C. This was followed by an extension at 60 °C. PCR products were electrophoresed on an ABI 3730 and alleles were scored using GeneMapper (ABI). *Omm1046* and *Ots100* were not used for further analysis due to poor amplification.

The spawning time gene *OtsClock1b* was also amplified following the conditions in O'Malley (2008a). The fluorescently labelled sense primer was 5'-CCTGTGTTTGTCTCCAACAGCA-3' and the antisense primer was 5'-CTGTCAGTGC-AAATTACAGTCCT-3'. Genomic DNA was amplified in 10 µl reactions using a touchdown PCR profile. This consisted of one initial denaturing cycle of 3 min at 94°, followed by one cycle of 1 min at 94 °C, 1 min at 62 °C annealing temperature and 1 min 30 s at 72 °C. In subsequent cycles, the annealing temperature was decreased by 2° until 56 °C was reached, followed by 29 more cycles of 1 min 94 °C, 1 min at 56 °C, 1 min 30 s at 72 °C and a final extension of 10 min at 72 °C. After electrophoresis on an ABI 3730, polymorphisms were

scored using GENEMAPPER software (ABI). Raw microsatellite genotype data was binned into size ranges for allele calls using the program *tandem* (Matschiner and Salzburger 2009).

Table 5.1. Microsatellite loci, fluorescent dye label used for visualization, and PCR concentration in a 5 μ l reaction for seven markers amplified in this study.

| locus | dye label | PCR concentration (mM) |
|----------------------|-----------|------------------------|
| <i>Omm1046</i> | NED | 0.15 |
| <i>Ots100</i> | FAM | 0.30 |
| <i>Ssa407</i> | PET | 0.40 |
| <i>Omy1011</i> | VIC | 0.16 |
| <i>Omy1001</i> | NED | 0.16 |
| <i>Ots3m</i> | NED | 0.20 |
| <i>OmyFGT12TUF-F</i> | PET | 0.20 |

Statistical analysis

Conformance to Hardy-Weinberg equilibrium (HWE) and the presence of gametic disequilibrium were evaluated using GENEPOP on the web (Raymond and Rousset 1995). Locus-by-locus pair-wise tests for gametic disequilibrium utilized 1000 dememorizations, 100 batches, and 1000 iterations per batch. Comparisons were considered significant when $P < 0.05$ with a sequential Bonferroni correction (Holm 1979). Hardy-Weinberg exact tests were performed for each locus in the population using a probability test, with 1000 dememorizations, 100 batches, and 1000 iterations per batch.

Pairwise F_{IS} , F_{ST} , and expected heterozygosity (H_E) were calculated using FSTAT 2.9.3 (Goudet 1995). Allelic richness was calculated using HP-RARE v.1.0 (Kalinowski 2005), which uses a rarefaction method to correct for differing sample sizes, with the smallest sample size used for calculation being $N = 16$. A principal components analysis (PCA) was performed in PCAGEN (J. Goudet unpublished; <http://www.unil.ch/izea/software/pca-gen.html>) in order to examine the variation between populations.

A power analysis of sample size was performed using the spatial genetic dataset from Chapter 3 using data from the seven markers used in this study. Pairwise F_{ST} was calculated using GENEPOP on the web (Raymond and Rousset 1995) between Hinemaiaia and nine other populations with 37, 27, and 17 individuals, with individuals randomly removed from the

Hinemaiaia population to create the smaller test populations. The percentage difference between the largest sample set and the smaller samples was calculated as the difference in F_{ST} divided by the average across markers. This process was repeated for Waimarino.

5.3 Results

Genetic diversity

In the analysis of historic samples there was no evidence of contamination in any of the negative controls. The number of alleles per locus ranged from 3 to 14. 18 private alleles were observed, the majority of them (15) in the 2006 Hinemaiaia sample (Table 5.2). All populations had similar allelic richness and expected heterozygosity. F_{IS} was lowest in the 2008 Waimarino sample (-0.045), and highest in the 1987 Hinemaiaia sample (0.176). The highest F_{IS} value for Hinemaiaia is likely attributable to high F_{IS} at the *Omy1001* locus, which occurs in both Hinemaiaia samples and also in the 1987 Waimarino sample. Samples not in HWE were the Hinemaiaia 2006 at *Omy1011* and *Omy1001*, Hinemaiaia 1987 at *OmyFGT12TUF*. MICROCHECKER indicated that null alleles may be present at *Omy1011* in Tokaanu 1987 and Hinemaiaia 2006 and *Omy1001* in Hinemaiaia 2006, and *OmyFGT12TUF* in Hinemaiaia 1987 due to a homozygote excess in these populations. I observed no gametic disequilibrium after a strict Bonferroni correction.

Table 5.2. Sample information, including river, sampling year, and number of samples (N), and summary statistics, including F_{IS} , allelic richness (Ar), expected heterozygosity (H_E), and the number of private alleles (PA).

| River (abbreviation) | Time of sampling | N | Source of DNA | F_{IS} | Ar | H_E | PA |
|-------------------------|---------------------|-----|------------------|----------|------|-------|----|
| Hinemaiaia(H) | Sept 1987 | 19 | scales | 0.176 | 5.35 | 0.759 | 1 |
| | Aug 2006 | 37 | tissue | 0.153 | 6.12 | 0.795 | 15 |
| Tokaanu (T) | Sept 1987 | 14 | scales | 0.142 | 5.10 | 0.796 | 0 |
| Waimarino(W) | Sept 1987 | 24 | scales | 0.100 | 5.55 | 0.784 | 1 |
| | Oct 2008 | 32 | tissue | -0.045 | 5.17 | 0.770 | 1 |

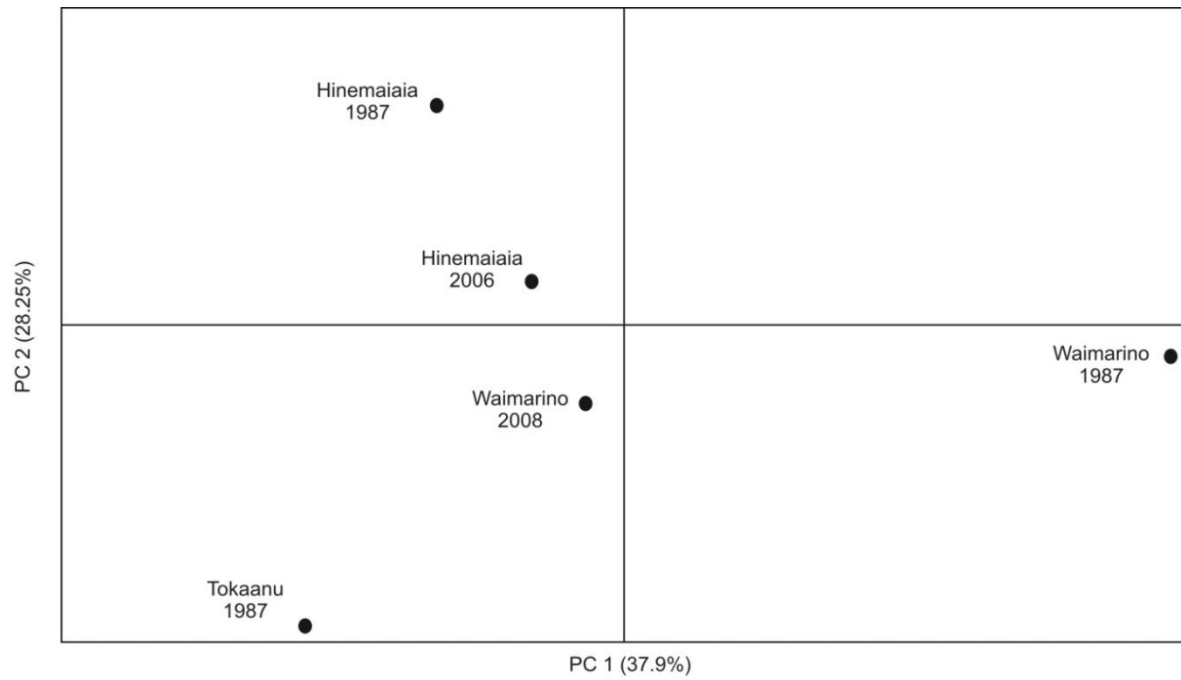


Figure 5.1. A principal components analysis (PCA) of temporal samples from three Lake Taupo populations.

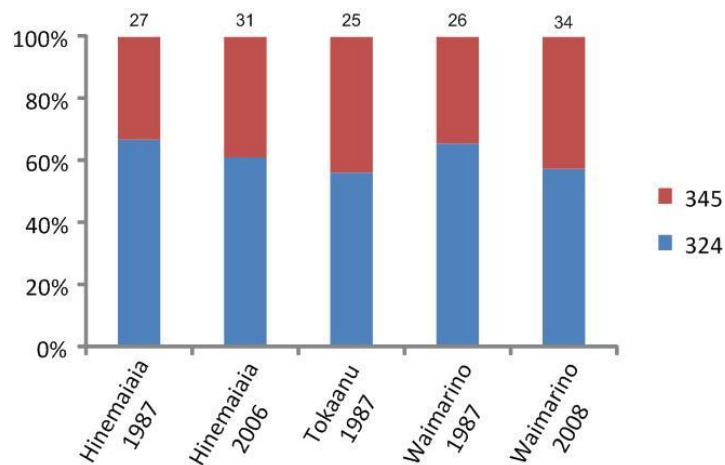


Figure 5.2. Allele frequencies of three Lake Taupo populations for the poly-Q region of *OtsClock1b* over time. Sample sizes are at the top of the bars. There is no significant difference in the frequency between populations ($\chi^2 < 2.4$, $P > 0.05$) as tested by genic differentiation.

Table 5.3. Population genetics statistics in scale samples compared to recent samples, including F_{IS} , allelic richness (Ar), expected heterozygosity (H_E), the number of private alleles (PA), and the total number of alleles (N_A). Sample abbreviations are the first letter of the sample site and the collection year (see Table 5.2 for full names).

| Sample | OmyFGT12_TUF | | | | | | Omy1011 | | | | | | Omy1001 | | | | | | Ssa407 | | | | | | Ots3M | | | | | |
|--------|----------------|----|------|----------------|-----------------|--|----------------|----|------|----------------|-----------------|--|----------------|----|------|----------------|-----------------|--|----------------|----|------|----------------|-----------------|--|----------------|----|------|----------------|-----------------|--|
| | N _A | PA | Ar | H _E | F _{IS} | | N _A | PA | Ar | H _E | F _{IS} | | N _A | PA | Ar | H _E | F _{IS} | | N _A | PA | Ar | H _E | F _{IS} | | N _A | PA | Ar | H _E | F _{IS} | |
| H1987 | 4 | 1 | 3.49 | 0.632 | 0.583 | | 7 | 0 | 6.18 | 0.841 | 0.027 | | 7 | 0 | 5.90 | 0.756 | 0.186 | | 9 | 0 | 7.29 | 0.870 | 0.092 | | 4 | 0 | 3.91 | 0.698 | 0.079 | |
| H2006 | 5 | 2 | 3.61 | 0.621 | 0.086 | | 12 | 3 | 7.26 | 0.863 | 0.217 | | 14 | 6 | 7.37 | 0.845 | 0.265 | | 9 | 0 | 7.39 | 0.874 | 0.024 | | 8 | 4 | 4.95 | 0.773 | 0.161 | |
| T1987 | 3 | 0 | 3.00 | 0.663 | 0.304 | | 6 | 0 | 5.45 | 0.841 | 0.459 | | 7 | 0 | 6.56 | 0.861 | -0.045 | | 7 | 0 | 6.51 | 0.872 | 0.029 | | 4 | 0 | 4.00 | 0.741 | -0.012 | |
| W1987 | 3 | 0 | 2.96 | 0.630 | 0.207 | | 8 | 0 | 6.68 | 0.857 | 0.056 | | 8 | 0 | 6.36 | 0.822 | 0.189 | | 11 | 1 | 7.98 | 0.893 | -0.026 | | 4 | 0 | 3.75 | 0.718 | 0.113 | |
| W2008 | 4 | 1 | 3.21 | 0.641 | -0.171 | | 8 | 0 | 6.06 | 0.831 | 0.097 | | 7 | 0 | 5.68 | 0.801 | -0.053 | | 9 | 0 | 7.01 | 0.853 | -0.051 | | 4 | 0 | 3.91 | 0.723 | -0.081 | |

Table 5.4. Pairwise estimates of genetic differentiation (F_{ST}) from 5 microsatellite loci. F_{ST} values are below the diagonal and P values above the diagonal, obtained after 200 permutations. The adjusted significance level after a Bonferroni correction for multiple comparisons was $P < 0.005$. See Table 1 for full names of sampling locations).

| | H1987 | H2006 | T1987 | W1987 | W2008 |
|-------|-------|-------|-------|-------|-------|
| H1987 | -- | 0.160 | 0.190 | 0.600 | 0.280 |
| H2006 | 0.000 | -- | 0.220 | 0.550 | 0.060 |
| T1987 | 0.000 | 0.000 | -- | 0.140 | 0.360 |
| W1987 | 0.007 | 0.006 | 0.012 | -- | 0.200 |
| W2008 | 0.001 | 0.000 | 0.000 | 0.007 | -- |

Genetic differentiation

Comparison of pairwise estimates of genetic differentiation (F_{ST}) indicates that there was no significant genetic divergence between populations (Table 5.4). F_{ST} values between populations ranged from 0.000 to 0.012, with a global F_{ST} of 0.022. The PCA showed that both samples from the Hinemaiaia clustered together in principle component PC 1, while the Waimarino samples were most similar on PC 2 (Figure 5.1). The Tokaanu sample was most similar to the Hinemaiaia samples.

Allele frequencies at *OtsClock1b* have remained constant over time (Figure 5.2) and there is no statistical difference in the frequency of the 345 bp allele between samples, as determined with an exact test of genic differentiation ($\chi^2 < 2.4$, $P > 0.05$).

The power analysis indicated that a decrease in sample size could substantially change the pairwise F_{ST} values calculated, particularly if the population contains less than 15 individuals (Figure 5.3).

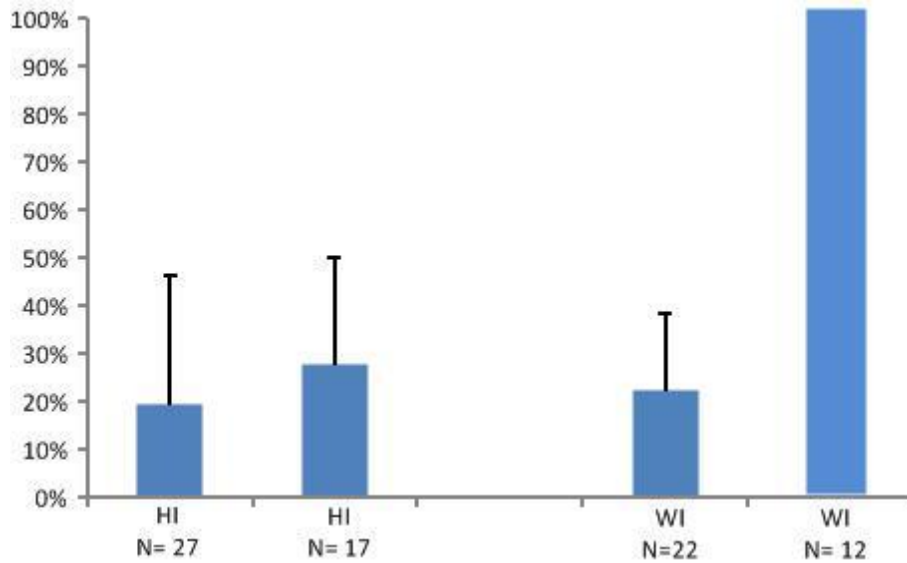


Figure 5.3. Power analysis of the effect of sample size for Hinemaiaia and Waimarino populations when compared with their true sample sizes ($N = 37$ and $N = 32$ respectively) for the modern day sample used in this study. Bars represent positive standard deviation. The Waimarino $N=12$ sample had a percentage difference of 246% with a standard deviation of 5.0.

5.4 Discussion

The Lake Taupo population shows possible genetic stability over time following introduction. Lake Taupo has a greater proportion of the 345 bp allele than any of the Californian founding populations, indicating either a founder effect or selection over time for this large allele. But there is no detectable change over between populations in 1987 and 2006/2008 in allele frequencies at *OtsClock1b* in two Lake Taupo populations. Some exploited populations see a decline in genetic diversity, particularly heterozygosity, over time (Miller and Kapuscinski 1997), but no similar trend has been detected in Lake Taupo.

Genetic diversity

The 2006 Hinemaiaia sample did have a much larger number of private alleles than the Hinemaiaia 1987 sample, but these alleles were at such low frequency that they did not result in a significant difference between the samples. Allelic diversity is often more sensitive than H_E for detecting changes in population size, so could be a better indicator for fluctuations in the population (Schwartz *et al.* 2007), and therefore this may be an indication

of recent demographic and/or migratory changes to the population. These low frequency alleles could be a sign of recent migration into the Hinemaiaia population from elsewhere in the lake, and if they increase in frequency would lead to changes in the genetic differentiation of the Hinemaiaia population. Alternately, these low frequency alleles might not have been observed in the historic population because of poor amplification of the older, possibly more degraded samples. If poor amplification had been a factor, however, there would be similar results for the Waiotaka.

The historic Waiotaka and Hinemaiaia samples had only one private allele each, indicating that only one allele from each population has been lost over the last twenty years. Loss of rare alleles is one of the first signs of a loss of genetic diversity, and in New Zealand snapper indicated that population size had become small enough for genetic drift to have a large effect in the population (Hauser *et al.* 2002). Lake Taupo rainbow trout do not appear to have undergone a loss of genetic diversity despite their continued exploitation, indicating that effective population size is likely high enough that genetic drift does not have a large effect on populations. Since there was no evidence of strong genetic drift, it can be concluded that the higher frequency of the 345 bp allele at *OtsClock1b* in Lake Taupo populations compared to California source populations is due to either a founder effect or selection.

Californian populations either did not have the 345 bp allele at *OtsClock1b* or it was present at low levels (6% or less), as compared to a minimum of 26% in one temporal sample in Lake Taupo, and on average was present at 29.8% in a random sample of Taupo individuals (Chapter 4). While this difference in allele frequency could be due either to a founder effect or to selection, there was no significant variation in *OtsClock1b* allele frequencies observed between samples from the 1980s and the 2000s. It is therefore unlikely that the frequency of the 345 bp allele has increased recently, but it is not clear from these results whether the frequency has been higher in Taupo than California since introduction, or if there was possibly a selective sweep early in the history of Taupo rainbow trout that caused them to have a higher frequency of the large allele.

Genetic differentiation

Pairwise F_{ST} comparisons revealed that there is no significant variation between current samples and samples from the 1980s. The power analysis performed on sample size,

though, indicates that the low sample sizes utilized for the scale analysis could substantially affect the calculation of pairwise F_{ST} , particularly when sample size is lower than 15 individuals (Figure 5.3).

In spite of the confounding effects of sample size, the finding that there has not been substantial change in the population differentiation of Hinemaiaia and Waimarino populations fits a general pattern that is often observed in multiple historic population genetics studies where, in the absence of a major catastrophic event, populations will maintain much of their genetic structure over time (Tessier and Bernatchez 1999; Ruzzante *et al.* 2001; Hansen 2002; Brown *et al.* 2007). It is important to test for stability in temporal patterns of genetic differentiation since it is an underlying assumption of many population genetic analyses (Waples 1998). In the case of British Columbia steelhead a change in genetic structure, as well as a decrease in genetic diversity and lowered effective population size over time was not anticipated, but was hypothesized to be caused by a landslide event in the studied catchment (Heath *et al.* 2002). This result emphasizes the importance of analysing historical samples whenever they are available to gain a more complete picture of the underlying causes of population structure.

Conclusions and future work

The relative genetic stability that was observed in Lake Taupo should be viewed cautiously in light of the limited sample size and low number of genetic markers. Due to a limited numbers of individuals and polymorphic loci estimates of F -statistics could have been underestimated (Waples 1989). Sufficient sampling is important in historical studies because historical data are error prone (Wandeler *et al.* 2007). To better understand the historical population dynamics of Taupo rainbow trout this study would benefit from a current-day Tokaanu sample and the analysis of further individuals and loci from both historic and current sampling groups. The inclusion of a modern day Tokaanu sample is currently being undertaken and will be included in any future publications.

Luikart *et al.* (1998b) recommends that one should sample a minimum of five highly variable loci from 30 individuals to have a greater than 80% chance of detecting a bottleneck with an effective size of 10 individuals, and therefore a significant loss of heterozygosity, after one generation. This study examines individuals with more than one generation

between sampling points and this dataset barely meets the criteria for a high probability of detecting changes in allele frequency. To make definitive conclusions about the temporal genetic variation in Lake Taupo populations, particularly in light of the change in the number of private alleles at Hinemaiaia, it will be necessary to sample more loci.

Overall these results are positive for the continued viability of Lake Taupo populations. Temporal variation should be a major consideration when interpreting population genetic structure (Heath *et al.* 2002). Allele frequency can be expected to vary year to year through genetic drift alone (Waples and Teel 1990), but variations in effective population size can have large effects on genetic structure, and it is important to analyse historic samples to understand this variation (Heath *et al.* 2002).

CHAPTER 6

General Discussion

6.1 Introduction

This study of population genetic structure and spawn timing genetics of Lake Taupo rainbow trout contributes to a broader understanding of the effects of translocation and the genetics of salmonid spawning time. The aims of this study were to 1) determine the source populations for the New Zealand trout introduction, specifically the founding populations for Lake Taupo, 2) determine whether there was genetic differentiation among populations from tributaries around Lake Taupo, 3) determine whether there was a correlation between variation at the *OtsClock1b* gene and the spawn timing of rainbow trout, and 4) determine whether there were temporal changes in population genetic structure and allele frequencies at locus *OtsClock1b*.

In this chapter I will discuss the implications of my findings, and present management recommendations for the Lake Taupo fishery that arise from the genetic study. Initially this study was undertaken in collaboration with the Department of Conservation to establish a population genetic baseline for Lake Taupo and determine the causes for the decline of their early spawning run. Over the last few months a stocking program has been proposed as a method for trying to restore the early run of Lake Taupo rainbow trout by taking individuals from nearby manmade Lake Otomangakau (Target Taupo July 2011, Issue 63). The management recommendations in this chapter therefore contain not only recommendations pertaining to the genetic structure and spawning time genetics of Lake Taupo trout, but also some recommendations pertaining to this proposed and currently operational stocking program.

6.2 Major findings and their implications

Genetic structure

Lake Taupo rainbow trout most likely come from multiple sources based on the results of Chapter 2 and historical records. Although multi-source introductions are thought to enable a species to subvert the so-called “invasion paradox” (Allendorf and Lundquist 2003), a bottleneck is still detectable in Lake Taupo rainbow trout. These results indicate that the introduction bottleneck and/or the subsequent demographic fluctuations were still strong enough to leave a persistent signature of lowered genetic diversity on the Taupo population compared to source populations. The lower level of genetic diversity observed at the fifty Lake Taupo samples used in Chapter 2 was consistent with the results from all ten populations sampled for the full spatial population study (Chapter 3). This is unusual considering that a recent review of freshwater introductions showed that successful invasive species often do not exhibit the genetic signatures of bottlenecks (Roman and Darling 2007). This finding emphasizes the importance of genetic monitoring of translocated populations, since allelic diversity can be lost even when the founding population is large (Stockwell *et al.* 1996). However levels of molecular diversity are not necessarily a predictor of ecological success, because variation at only a few particular genes could have a large impact on colonising ability (Lee 2002). The persistence and abundance of Lake Taupo trout despite their bottleneck signature exemplifies the success of an introduced species in spite of reduced genetic diversity.

The bottleneck simulations showed that Taupo trout are likely derived from both inland and coastal sources. If introduced individuals come from two distinct genetic groups (in this case inland and coastal Californian populations), the simulations indicated that detecting the true sources becomes difficult, if not impossible. This finding could be particularly important for controlling and managing introductions, since an inability to determine the true sources or signal pollution from an unknown source may make the results of origins studies inconclusive. Tracing introductions, also known as “routes of invasion”, is increasingly recognized as an important field of study if non-native translocations, both intentional and not, are to be properly managed (Estoup and Guillemaud 2010).

The analysis of spatial population structure in Chapter 3 showed that only populations whose migration is constrained by landscape barriers are differentiated in a small lake catchment, but there was not an overall pattern of genetic differentiation between populations within Lake Taupo. Dispersal barriers are known to have a large impact on increasing genetic structure, as they usually decrease gene flow (Wofford *et al.* 2005). This lack of gene flow therefore increases the possibility that genetic drift will cause a population to lose diversity (Slatkin 1985). Lake Tarawera in the Rotorua area did have the lowest allelic richness of all populations, which particularly makes sense given its history of hatchery stocking, since fish cannot spawn naturally in the Lake Rotorua habitat.

In spite of the lack of genetic differentiation at F -statistics, the presence of rare alleles in most Lake Taupo populations, even though they are at low frequency, indicates that there may be some locally developing differentiation. While rare alleles can be a sign of migration in long established population, in recently introduced populations, Slatkin's "radiation model" suggests that rare alleles merely represent a founder effect (Slatkin 1985). Over time these rare alleles become a true sign of migration within a population, because any rare alleles which resulted from a founder effect would become either more frequent in the population or be eliminated, due to genetic drift. In Chapter 5, I reported that the Hinemaiaia sample from 2006 shows an abundance of private alleles which are not present in a sample from the same site in 1986, which following Slatkin's radiation model would suggest that these alleles are likely present due to migration.

There is also the risk that the genetic differences observed between populations using pairwise F_{ST} values are actually signs of population subdivision, and that the Bonferroni correction used for multiple tests to prevent Type I error actually causes Type II error. While some other measures, including false discovery rate (FDR) (Benjamini and Hochberg 1995) and modified FDR (Narum 2006) show less Type I error than Bonferroni, there is still quantitative way to evaluate whether statistical significance correlates with biological significance (Waples 1998; Hedrick 1999; Balloux and Lugon-Moulin 2002; Narum 2006). Because multiple analyses, including the Bayesian structure analysis and the FCA, indicate that Taupo populations are not structured, using the more conservative Bonferroni correction is likely appropriate, considering that most of the analyses performed in Chapter 3 indicate that there is more danger with this dataset of Type I error occurring than Type II, or that

structure would be inferred where none is present. Throughout this dissertation the P value is also usually accompanied by the test statistic, since the reporting of the effect size, and allowing the reader to evaluate the meaningfulness of the subsequent P value, is another way to combat the Type II error introduced by Bonferroni corrections (Nakagawa 2004). This illustrates the importance of utilizing multiple analyses, not just pairwise F -statistics, to evaluate population genetic structure so that results are biologically relevant and not only statistically significant.

There is no evidence of any recent population turnover in Lake Taupo populations, as there have been no relevant local catastrophic events that significantly affected the lake rainbow trout. It is possible that there has been sufficient exploitation pressure, or migration in the lake itself to cause a shift in genetic differentiation and/or diversity over the last twenty years. Indeed this study determined that while rare alleles in the present Hinemaiaia sample may indicate migration within the lake, there is not further evidence of temporal instability in the two populations examined here. This result is in keeping with other studies of salmonid population structure which showed that even in disturbed conditions salmon maintain stable population structuring (Nielsen *et al.* 1997; Nielsen *et al.* 1999b; Tessier and Bernatchez 1999). Future geological instability could cause genetic changes, as was the case with British Columbian steelhead, whose population structure was likely impacted by a landslide (Heath *et al.* 2002). This is a very real possibility for Lake Taupo trout considering that the lake is contained within a caldera on a volcanic plateau. There were volcanic eruptions of Mount Ruapehu, in close proximity to Lake Taupo, in 1995 which caused ash to fall on some of the tributaries around the lake, and since a tracking study was underway, some immediate changes in trout migration around the lake were observed, although trout returned to their previous spawning tributaries within two weeks (M. Dedual pers comm.). There was also some flooding which caused trout to augment their spawning habitat to spawn further upstream, although migration to different spawning tributaries was not indicated (Venman and Dedual 2005). It is possible that events of larger magnitude could lead to migration between tributaries or even fish kill in certain populations and therefore lead to changes within the demography and genetic structure of the lake.

Spawn timing genetics

This study is the first to report two alleles in the poly-Q region of *OtsClock1b* locus amplified from rainbow trout, and the larger previously undiscovered allele was shown to be associated with later spawning time. This expands the growing volume of literature on the association of larger poly-Q regions at timing loci with later reproductive behaviours (Johnsen *et al.* 2007). A generalized linear model indicated that the large allele frequency varied with season in Lake Taupo populations, but not with fishing type or in relation to spatial location around the lake. This result makes sense in light of recent studies which show that day length at spawning accounts for much of the *OtsClock1b* allele frequency variation in two species of salmonid (O'Malley *et al.* 2010). However the study reported herein did not show a strong correlation between allele frequency and season at all sites, which may be due to environmental variation. While photoperiod is a major cue for migration for spawning (Quinn and Adams 1996), these processes may be delayed if certain physiological thresholds are not met (Arnesen *et al.* 2003). Environmental plasticity is one of the characteristics which make spawning migration time a difficult trait to study.

There are limitations to the candidate gene approach, since this study only examined one portion of one gene involved in photoperiodic behaviours. A recently published study, which utilized 92 SNP markers linked to 75 functional genes tested for a correlation with run-timing, found three loci which may be important factors in the phenotype of upriver migration time (Hess and Narum 2011). This sort of genomic scan approach might have greater utility in elucidating the effects that many different sites throughout the genome have on migration and spawning time. However, the lack of specific run-timing information for fish of known origin is a shared limitation of Hess and Narum (2011) and the study discussed herein. Without knowing the natal location of the fish being studied it is difficult to conclude whether straying migrants are being examined or true migrating individuals who seek to spawn upstream. There is a great need in this field for studies with adequate phenotypic data and/or greater control over environmental conditions. In the future spawning time genetics studies would greatly benefit from a combination of genomic scans and precise phenotypic spawning data.

Future work

It is important to note that although season was the only significant factor in the generalized linear model, the data were overdispersed, which indicates that there is variance within the data that the factors tested did not address. Temperature is known to be a major factor in the timing of salmonid migration (Heggberget 1988; Quinn and Adams 1996; Quinn 2005). River discharge has also been suggested as an important factor for salmonid migration, with fish preferring to move upstream after periods of high flow (Trépanier *et al.* 1996). Therefore it would be advisable to collect and analyse further environmental data to determine what the other significant factors are that determine allele frequency at *OtsClock1b*.

6.3 Management recommendations

Spatial genetic management

The current genetic structure of rainbow trout supports the management of Lake Taupo as a single fishery stock. However, it would still be reasonable to exercise the precautionary principle and manage the fishery by tributary in case genetic differentiation is in the process of forming. Freshwater fish species typically form genetically structured populations (Gyllensten 1985). It is important in setting management units not to just reject or fail to reject panmixia, but to incorporate the genetic diversity within a managed system into the delineation of management areas and population groups (Palsbøll *et al.* 2007).

I did find some evidence of outlier loci and private population alleles, which suggests the possibility of developing local differentiation or adaptation. If re-stocking was to occur in Lake Taupo there is no compelling evidence, based on the data presented here, to suggest that any particular local area having an unimpeded migration pathway to the lake should not be used as a source. If any populations that are currently isolated from Lake Taupo were to be source populations for stock movement (e.g. the Te Whaiau trap at Lake Otamangakau) then it would be advisable to conduct a specific genetic study to determine the level of differentiation, since the two populations of the present study (Rotorua and Waipakihi) that are physically isolated from Lake Taupo were the only populations that were found to be significantly different. If considering restocking from Rotorua or Waipakihi, the genetic

differentiation of these two populations could negatively impact the local population and/or the success of the introduced trout. Further investigation of differentiation at loci potentially under selection (e.g. *OtsClock1b*) and modelling of the effects of stocking would help to determine the extent of the risk involved with restocking from these geographically isolated populations. It would also be advisable to investigate genes associated with disease resistance (i.e. major histocompatibility complexes) in case there were a disease event in the lake, since a lack of genetic diversity at these loci could be detrimental were such an event to occur.

Recovery from catastrophic events can occur naturally, with recolonization by surviving individuals (Lamberti *et al.* 1991; Swanson *et al.* 1998). Radio telemetry studies of the Taupo catchment have shown that rainbow trout have augmented their spawning and migration sites after flooding events within the fishery (Venman and Dedual 2005), and it is likely they would do the same in the face of larger events. Waiting for natural recovery to occur may be the best management recommendation for Lake Taupo rainbow trout, particularly since stocked fish are known to be more susceptible to angling than wild fish (Dwyer and Piper 1984), and many fisheries management programs in Europe are now implementing native population supplementation programs (Caudron *et al.* 2009).

Seasonal variation: restoring the early run of Taupo rainbow trout?

The high frequency of the 345 bp allele in the Lake Taupo population compared to the lower frequencies found in the Californian source populations was most likely the consequence of the random choice of founding fish at the time of introduction. However, the process of selection (either natural or due to harvesting pressure) could also explain this high frequency pattern. Genetic variation at the *OtsClock1b* locus and the seasonal catches of anglers should be investigated. The seasonal *OtsClock1b* pattern at Waipa trap and the somewhat temporal stability of allele frequencies since the 1980s suggests that it may be possible to select for early spawners for restocking purposes. A plan for the supplementation of early spawners has already been proposed (Target Taupo 63) and it has been suggested that individuals could be introduced to Lake Taupo from the Te Whaiiau trap at Lake Otamangakau. A specific study should be conducted to determine the level of variation at microsatellite loci and the *OtsClock1b* locus because the possibility of genetic differentiation has been suggested in a previous study. The proposed stocking program may need to be

augmented (i.e. fewer fish released or fish stocked from a different location) to ensure minimal negative impact on the Tongariro fishery.

An important component of developing a restocking program will be assessing the potential impacts of introducing hatchery-reared broodstock on the genetic structure of existing populations. Therefore it would be prudent to investigate this issue using simulated models of the possible genetic impacts of restocking under difference scenarios (e.g. size of the restocking populations and its genetic composition). For example, computer models could simulate the introduction of Te Whaiau and/or Rotorua genotypes into the Tongariro catchment and the possible consequences of genetic drift. This work would be similar to the genetic bottleneck simulations performed as part of the origins work in this present research project, and would allow managers to design a restocking program the minimised any possible risks of genetic change to the existing populations.

It would be useful to compare a third modern sample to the historic samples, so that an investigation could be conducted into whether allele frequencies have changed in the Tokaanu population. This sample was recently collected and would add much to the study of temporal genetic variation in Lake Taupo populations. It would also be useful to add more loci to the comparison of current and historic population structure. More data would be useful in determining an estimate of migration rate between populations in Lake Taupo, which would in turn allow for a more quantitative determination of the appropriateness of managing the lake as one, or conversely many, stocks, depending on the migration rate. Optimally these genetically determined migration rates would be related back to the demographic independence of the target population to yield a management program which fully reflected the demography of Lake Taupo trout.

6.4 Contribution to the field

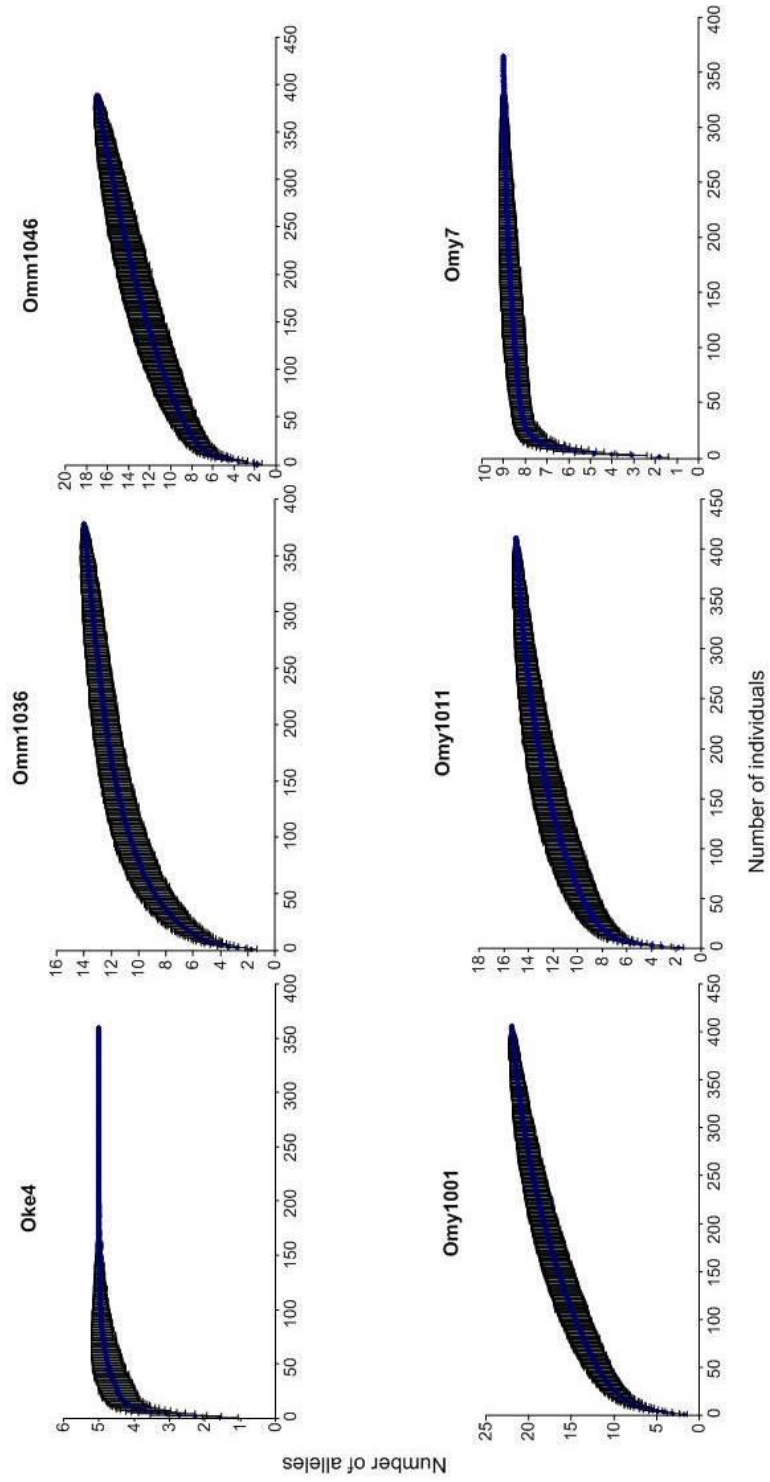
This study of introduced rainbow trout makes several important contributions to introduction ecology and fisheries management. The simulations of introduction scenarios performed for Chapter 2 show that multi-source introduction scenarios are difficult to interpret, but simulations can help identify the most likely true history. The study of introduced species which are endangered and threatened in their native range offers a unique

opportunity to inform conservation in a somewhat circuitous way. My finding that seasonal spawning migration might have an association to allele frequency at the *OtsClock1b* gene may have commercial applications for domesticated trout strains. It will also inform conservation management of native *O. mykiss*. Understanding translocated salmon populations will be particularly important as new habitats emerge, due to opening of new stream habitats through climate change or the addition of fish passages to dams, or as current salmon stocks invade new areas, as has happened in Patagonia (Becker *et al.* 2007). It's possible that the genotypic differences in spawning time that were observed in Waipa may eventually lead to differentiation at neutral markers as well, and there by the segregation of different spawning times into separate spawning populations. Quinn (2000) have shown that divergence in spawning time can drive divergence at other traits, and it is likely that Taupo rainbow trout could eventually show the same divergence as South Island Chinook populations.

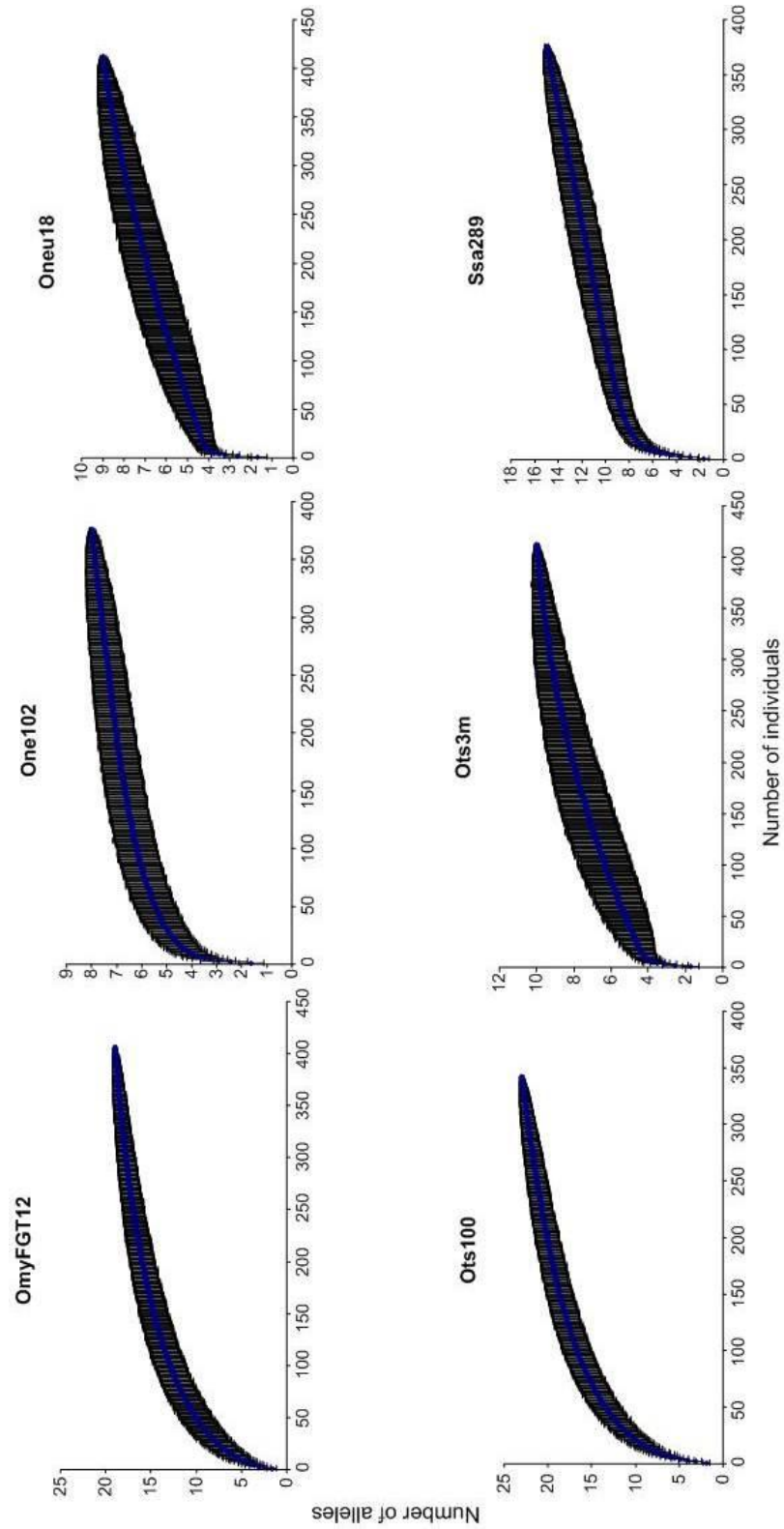
Salmonids phenotypic plasticity and duplicated genome make them a challenge suite of species to study genetically, but also make them a fascinating candidate for studies of contemporary evolution and the response to environmental change. While I suspect that an environmental change has caused the depletion of the early run of fish in Lake Taupo, and it may never be known what that environmental trigger was, that depletion provided an opportunity to investigate the population genetics of a New Zealand rainbow trout population more thoroughly than they have ever been examined before. Hopefully this is only the beginning of more intensive genetic studies of introduced salmonid populations, and the elucidation of the influence that circadian oscillators have on salmonid spawning migration.

Appendix

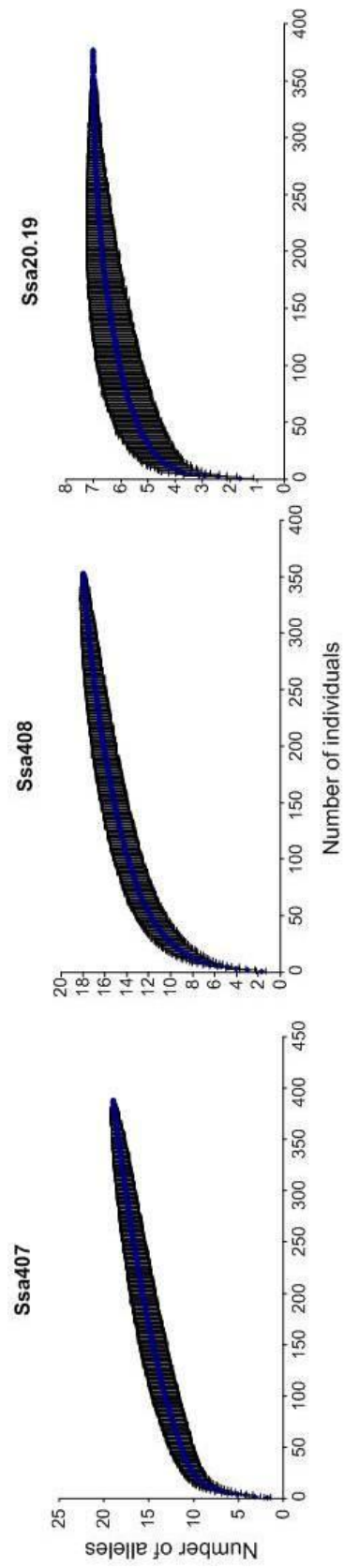
App. Figure 3.1.1. Allele discovery curves for ten Lake Taupo and one Rotorua population, computed by jackknife simulation.



App. Figure 3.1. (cont)

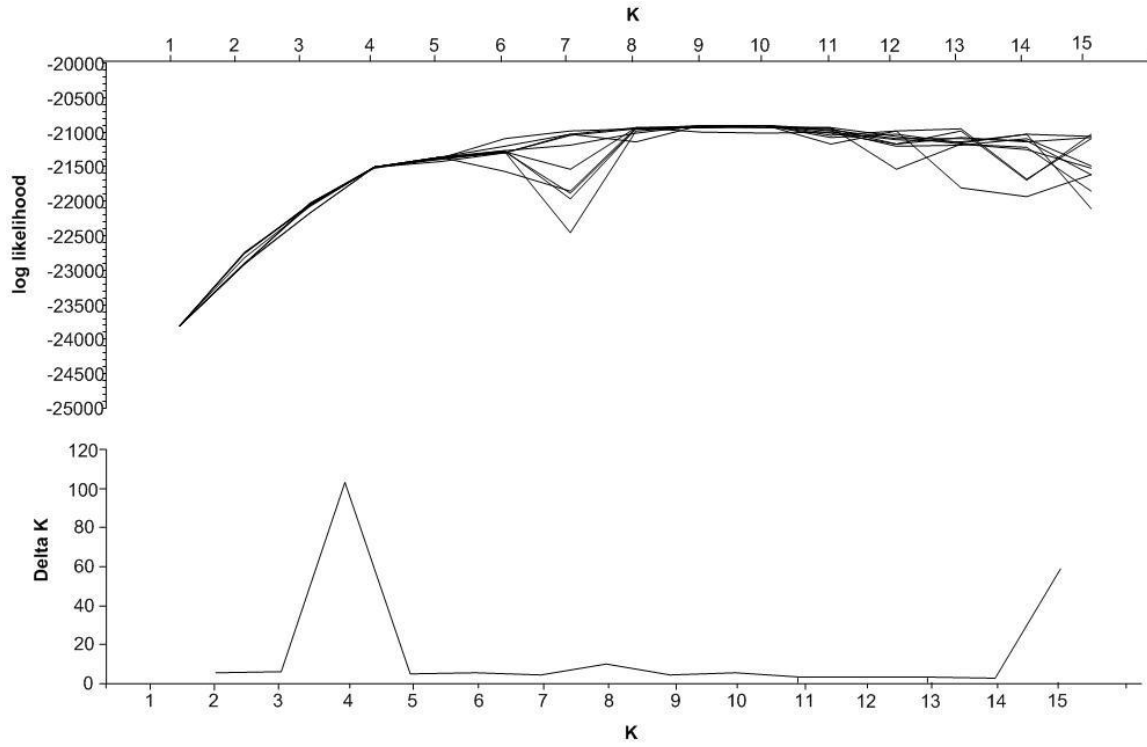


App. Figure 3.1. (cont)



Appendix

App. Figure 3.2. Determination of number of K cluster from population genetic data from ten Lake Taupo populations and one Rotorua population. The upper graph is the log likelihood for values of K from 1 to 15, while the lower is the (Evanno *et al.* 2005) delta K value for values of K from 2 through 15. The largest average log likelihood is K=10 (-20775.6).



App. Table 3.1. Population genetic statistics for ten Lake Taupo populations and one Rotorua rainbow trout population.
 PA = private alleles, N_A = total number of alleles.

| | <i>Omy1001</i> | | | | | | <i>Omm1046</i> | | | | | | <i>Omm1036</i> | | | | | | <i>Oke4</i> | | | | | |
|---------------|----------------|------|------|-------|----------|--|----------------|------|------|-------|----------|--|----------------|------|------|-------|----------|--|-------------|------|------|-------|----------|--|
| | N_A | PA | Ar | H_E | F_{IS} | | N_A | PA | Ar | H_E | F_{IS} | | N_A | PA | Ar | H_E | F_{IS} | | N_A | PA | Ar | H_E | F_{IS} | |
| Hinemaiaia | 14 | 1 | 6.36 | 0.845 | 0.265 | | 8 | 0 | 5.11 | 0.778 | -0.077 | | 9 | 0 | 4.40 | 0.716 | 0.057 | | 4 | 1 | 3.52 | 0.593 | -0.097 | |
| Kuratua | 12 | 0 | 6.27 | 0.853 | -0.096 | | 8 | 1 | 5.20 | 0.800 | -0.141 | | 9 | 0 | 4.57 | 0.729 | -0.097 | | 4 | 0 | 3.35 | 0.572 | 0.046 | |
| TT | 11 | 0 | 5.51 | 0.809 | -0.047 | | 8 | 0 | 4.75 | 0.733 | -0.080 | | 6 | 0 | 4.42 | 0.766 | 0.158 | | 4 | 0 | 3.56 | 0.639 | -0.017 | |
| Waihukahuka | 9 | 0 | 5.71 | 0.815 | -0.018 | | 6 | 0 | 4.56 | 0.742 | -0.034 | | 8 | 0 | 3.96 | 0.695 | -0.064 | | 4 | 0 | 3.34 | 0.567 | 0.026 | |
| Waimarino | 7 | 0 | 5.17 | 0.801 | -0.053 | | 9 | 0 | 4.95 | 0.764 | -0.135 | | 9 | 0 | 4.85 | 0.783 | 0.042 | | 4 | 1 | 3.33 | 0.524 | -0.146 | |
| Waiotaka | 9 | 0 | 5.47 | 0.764 | 0.079 | | 10 | 2 | 5.03 | 0.779 | 0.059 | | 7 | 0 | 4.76 | 0.750 | -0.053 | | 4 | 0 | 3.43 | 0.549 | 0.123 | |
| Waipa | 12 | 1 | 5.64 | 0.821 | 0.050 | | 8 | 1 | 4.62 | 0.684 | 0.005 | | 7 | 0 | 4.28 | 0.761 | -0.051 | | 4 | 0 | 3.45 | 0.571 | -0.085 | |
| Waipakihi | 9 | 1 | 5.95 | 0.842 | -0.063 | | 6 | 1 | 3.84 | 0.656 | -0.042 | | 4 | 0 | 3.65 | 0.689 | -0.005 | | 4 | 0 | 2.96 | 0.466 | -0.016 | |
| Whangamata | 8 | 0 | 5.74 | 0.824 | -0.040 | | 6 | 0 | 4.74 | 0.757 | -0.091 | | 6 | 0 | 4.50 | 0.647 | -0.069 | | 5 | 0 | 2.89 | 0.392 | 0.188 | |
| Whitikau | 9 | 0 | 4.99 | 0.792 | 0.070 | | 8 | 1 | 5.24 | 0.792 | 0.008 | | 7 | 0 | 4.28 | 0.751 | -0.103 | | 5 | 1 | 3.77 | 0.603 | 0.006 | |
| Average/Total | 20 | 3 | 5.76 | 0.817 | 0.015 | | 17 | 6 | 4.93 | 0.749 | -0.053 | | 14 | 0 | 4.52 | 0.729 | -0.019 | | 5 | 3 | 3.41 | 0.548 | 0.003 | |
| Rotorua | 12 | 2 | 7.04 | 0.891 | 0.008 | | 7 | 0 | 4.74 | 0.743 | -0.077 | | 6 | 0 | 4.31 | 0.732 | 0.015 | | 5 | 0 | 4.05 | 0.694 | -0.017 | |

App. Table 3.1.continued

| | Ots3m | | | | | Ots100 | | | | | Oneul8 | | | | | One14 | | | | |
|---------------|----------------|----|------|----------------|-----------------|----------------|----|------|----------------|-----------------|----------------|----|------|----------------|-----------------|----------------|----|------|----------------|-----------------|
| | N _A | PA | Ar | H _E | F _{IS} | N _A | PA | Ar | H _E | F _{IS} | N _A | PA | Ar | H _E | F _{IS} | N _A | PA | Ar | H _E | F _{IS} |
| Hinemaiaia | 8 | 2 | 4.57 | 0.773 | 0.161 | 9 | 0 | 4.42 | 0.732 | -0.324 | 4 | 0 | 3.81 | 0.738 | 0.051 | 4 | 0 | 2.96 | 0.614 | 0.743 |
| Kuratua | 5 | 0 | 3.70 | 0.680 | 0.042 | 11 | 1 | 5.97 | 0.839 | -0.047 | 4 | 0 | 3.87 | 0.747 | 0.118 | 6 | 0 | 3.43 | 0.692 | 0.407 |
| TT | 4 | 0 | 3.80 | 0.734 | 0.035 | 13 | 1 | 6.25 | 0.830 | -0.142 | 4 | 0 | 3.80 | 0.735 | -0.058 | 6 | 1 | 3.47 | 0.619 | 0.627 |
| Waihukahuka | 4 | 0 | 3.70 | 0.711 | 0.043 | 12 | 1 | 5.99 | 0.831 | -0.109 | 4 | 0 | 3.72 | 0.726 | 0.092 | 3 | 0 | 2.70 | 0.591 | 0.506 |
| Waimarino | 4 | 0 | 3.78 | 0.723 | -0.081 | 6 | 0 | 4.95 | 0.804 | -0.010 | 4 | 0 | 3.85 | 0.745 | -0.017 | 8 | 4 | 5.22 | 0.803 | 0.572 |
| Waiotaka | 4 | 0 | 3.71 | 0.713 | 0.098 | 8 | 1 | 5.87 | 0.854 | 0.122 | 4 | 0 | 3.72 | 0.718 | -0.057 | 4 | 0 | 3.03 | 0.595 | 0.664 |
| Waipa | 7 | 0 | 4.20 | 0.733 | 0.018 | 14 | 1 | 5.84 | 0.788 | -0.168 | 4 | 0 | 3.83 | 0.739 | -0.110 | 5 | 0 | 3.09 | 0.612 | 0.542 |
| Waipakihi | 4 | 0 | 3.48 | 0.569 | 0.075 | 5 | 0 | 3.60 | 0.645 | 0.483 | 4 | 0 | 3.45 | 0.668 | 0.028 | 4 | 0 | 3.01 | 0.617 | 0.573 |
| Whangamata | 4 | 0 | 3.77 | 0.733 | -0.008 | 7 | 0 | 5.02 | 0.764 | 0.019 | 4 | 0 | 3.76 | 0.727 | 0.284 | 3 | 0 | 3.00 | 0.733 | 0.773 |
| Whitikau | 5 | 1 | 3.95 | 0.745 | -0.068 | 11 | 0 | 6.29 | 0.852 | -0.035 | 7 | 2 | 4.42 | 0.770 | 0.171 | 5 | 0 | 3.72 | 0.670 | 0.751 |
| Average/Total | 10 | 3 | 3.91 | 0.711 | 0.032 | 21 | 5 | 6.18 | 0.794 | -0.021 | 7 | 2 | 3.85 | 0.731 | 0.050 | 12 | 5 | 3.86 | 0.655 | 0.616 |
| Rotorua | 4 | 0 | 3.63 | 0.718 | 0.158 | 7 | 2 | 4.16 | 0.716 | 0.132 | 7 | 2 | 3.99 | 0.713 | -0.012 | 3 | 0 | 2.93 | 0.630 | 0.419 |

App. Table 3.1.continued

| | <i>One102</i> | | | | | <i>OmyFGT12</i> | | | | | <i>Omy7</i> | | | | | <i>Omy1011</i> | | | | |
|---------------|---------------|----|------|-------|----------|-----------------|----|------|-------|----------|-------------|----|------|-------|----------|----------------|----|------|-------|----------|
| | N_A | PA | Ar | H_E | F_{IS} | N_A | PA | Ar | H_E | F_{IS} | N_A | PA | Ar | H_E | F_{IS} | N_A | PA | Ar | H_E | F_{IS} |
| Hinemaiaia | 4 | 0 | 3.56 | 0.700 | 0.047 | 5 | 1 | 3.30 | 0.616 | 0.066 | 8 | 0 | 5.94 | 0.837 | 0.064 | 12 | 1 | 6.40 | 0.863 | 0.217 |
| Kuratua | 5 | 1 | 3.71 | 0.705 | 0.211 | 4 | 0 | 2.49 | 0.533 | 0.105 | 7 | 0 | 5.06 | 0.798 | 0.009 | 10 | 0 | 6.02 | 0.853 | 0.108 |
| TT | 5 | 0 | 3.78 | 0.694 | -0.080 | 7 | 0 | 3.97 | 0.675 | 0.058 | 8 | 0 | 5.91 | 0.851 | 0.045 | 7 | 0 | 5.25 | 0.811 | -0.028 |
| Waihukahuka | 4 | 0 | 3.56 | 0.685 | 0.061 | 14 | 1 | 5.37 | 0.768 | -0.100 | 8 | 0 | 6.12 | 0.857 | 0.050 | 10 | 1 | 5.97 | 0.841 | -0.088 |
| Waimarino | 4 | 0 | 3.75 | 0.719 | 0.073 | 4 | 0 | 3.07 | 0.637 | -0.189 | 8 | 0 | 6.06 | 0.849 | 0.018 | 8 | 0 | 5.52 | 0.831 | 0.097 |
| Waiotaka | 4 | 0 | 3.47 | 0.661 | 0.433 | 4 | 0 | 3.49 | 0.638 | 0.104 | 8 | 0 | 6.07 | 0.846 | 0.081 | 8 | 0 | 5.71 | 0.840 | -0.014 |
| Waipa | 5 | 0 | 3.61 | 0.679 | 0.205 | 7 | 0 | 3.85 | 0.676 | 0.349 | 9 | 0 | 5.15 | 0.780 | -0.025 | 10 | 0 | 5.98 | 0.837 | 0.115 |
| Waipakihi | 4 | 0 | 3.92 | 0.769 | 0.247 | 6 | 1 | 4.71 | 0.770 | -0.230 | 8 | 0 | 5.33 | 0.794 | 0.280 | 6 | 0 | 5.16 | 0.817 | -0.159 |
| Whangamata | 4 | 0 | 3.73 | 0.701 | 0.069 | 8 | 0 | 4.51 | 0.704 | 0.148 | 6 | 0 | 5.16 | 0.817 | 0.054 | 7 | 0 | 5.41 | 0.841 | 0.173 |
| Whitikau | 5 | 1 | 3.56 | 0.662 | 0.137 | 11 | 1 | 5.58 | 0.789 | -0.122 | 8 | 0 | 5.30 | 0.770 | -0.160 | 10 | 0 | 6.22 | 0.864 | -0.038 |
| Average/Total | 8 | 2 | 3.79 | 0.698 | 0.140 | 17 | 4 | 4.48 | 0.681 | 0.019 | 9 | 0 | 5.80 | 0.820 | 0.042 | 15 | 2 | 5.91 | 0.840 | 0.038 |
| Rotorua | 3 | 0 | 2.78 | 0.480 | 0.059 | 4 | 2 | 2.3 | 0.5 | -0.11 | 7 | 0 | 5.27 | 0.8 | 0.026 | 6 | 0 | 4.73 | 0.793 | -0.086 |

App. Table 3.1.continued

| | Ssa20.19 | | | | | Ssa408 | | | | | Ssa407 | | | | | Ssa289 | | | | |
|---------------|----------|----|------|-------|----------|--------|----|------|-------|----------|--------|----|------|-------|----------|--------|----|------|-------|----------|
| | N_A | PA | Ar | H_E | F_{IS} | N_A | PA | Ar | H_E | F_{IS} | N_A | PA | Ar | H_E | F_{IS} | N_A | PA | Ar | H_E | F_{IS} |
| Hinemaiaia | 4 | 0 | 3.25 | 0.500 | 0.050 | 9 | 0 | 5.40 | 0.819 | 0.023 | 9 | 0 | 6.58 | 0.875 | 0.011 | 8 | 0 | 5.44 | 0.828 | 0.094 |
| Kuratua | 4 | 0 | 3.34 | 0.622 | -0.059 | 9 | 0 | 5.23 | 0.802 | -0.077 | 10 | 0 | 6.57 | 0.871 | 0.030 | 8 | 0 | 5.22 | 0.796 | 0.057 |
| TT | 5 | 0 | 3.47 | 0.641 | -0.006 | 9 | 0 | 5.73 | 0.835 | 0.182 | 10 | 0 | 6.21 | 0.856 | 0.075 | 8 | 0 | 4.90 | 0.763 | 0.095 |
| Waihukahuka | 4 | 0 | 3.64 | 0.669 | 0.110 | 9 | 0 | 6.31 | 0.874 | 0.101 | 10 | 0 | 6.49 | 0.868 | -0.071 | 8 | 0 | 5.56 | 0.807 | 0.071 |
| Waimarino | 5 | 0 | 3.76 | 0.676 | 0.148 | 13 | 0 | 6.66 | 0.875 | 0.115 | 9 | 0 | 6.23 | 0.855 | -0.014 | 12 | 5 | 6.54 | 0.854 | 0.233 |
| Waiotaka | 5 | 0 | 4.44 | 0.670 | 0.067 | 14 | 2 | 7.50 | 0.903 | 0.446 | 15 | 5 | 7.17 | 0.877 | 0.568 | 9 | 1 | 4.76 | 0.684 | 0.194 |
| Waipa | 5 | 0 | 3.59 | 0.657 | -0.004 | 9 | 0 | 6.02 | 0.853 | 0.039 | 9 | 0 | 6.45 | 0.870 | -0.012 | 7 | 0 | 5.17 | 0.808 | 0.159 |
| Waipakihi | 4 | 0 | 3.48 | 0.538 | 0.070 | 11 | 1 | 7.48 | 0.910 | 0.206 | 11 | 2 | 7.19 | 0.905 | 0.448 | 7 | 1 | 5.25 | 0.797 | 0.273 |
| Whangamata | 4 | 0 | 3.37 | 0.586 | 0.102 | 7 | 0 | 4.79 | 0.727 | 0.083 | 8 | 0 | 5.54 | 0.815 | 0.052 | 7 | 0 | 5.19 | 0.802 | 0.133 |
| Whitikau | 6 | 0 | 3.70 | 0.627 | -0.020 | 8 | 0 | 5.53 | 0.813 | 0.038 | 9 | 0 | 6.54 | 0.874 | -0.066 | 6 | 0 | 5.08 | 0.791 | -0.077 |
| Average/Total | 7 | 0 | 3.60 | 0.619 | 0.046 | 18 | 3 | 6.35 | 0.841 | 0.116 | 19 | 7 | 6.69 | 0.867 | 0.102 | 15 | 7 | 5.45 | 0.793 | 0.123 |
| Rotorua | 6 | 0 | 4.32 | 0.751 | -0.114 | 8 | 0 | 5.01 | 0.758 | 0.246 | 9 | 0 | 5.82 | 0.774 | 0.003 | 7 | 0 | 5.24 | 0.782 | 0.251 |

App. Figure 4.1 R code for the stepwise backwards elimination generalized linear model.

```
> clock1b <- read.csv("D:/R-D drive/wd/clock1b.csv")
> view(clock1b)
> clock.df <- clock1b
> clock.df$season <- as.factor(clock.df$season)
> clock.df$trap <- as.factor(clock.df$trap)
> clock.df$year <- as.factor(clock.df$year)
> clock.df$has324 <- as.factor(clock.df$has324)
> clock.df$has345 <- as.factor(clock.df$has345)
> clock.df$has345 <- as.factor(clock.df$has345)
> levels(clock.df$has345) <- c("N", "Y")
> str(clock.df)
'data.frame': 698 obs. of 9 variables:
 $ sex      : Factor w/ 2 levels "F","M": 1 1 1 1 1 1 1 1 1 2 ...
 $ forkleng: int 450 510 550 455 470 435 530 450 450 430 ...
 $ month    : Factor w/ 8 levels "August","December",...: 1 1 1 1 1 1 1 1 1
1 1 ...
 $ year     : Factor w/ 5 levels "2006","2007",...: 1 1 1 1 1 1 1 1 1 1
1 ...
 $ has324   : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
 $ has345   : Factor w/ 2 levels "N","Y": 1 1 1 2 2 1 1 1 1 2 ...
 $ trap     : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
 $ season   : Factor w/ 3 levels "1","2","3": 2 2 2 2 2 2 2 2 2 2 ...
 $ genotype : Factor w/ 3 levels "0|1","1|0","1|1": 2 2 2 3 3 2 2 2 2 2
3 ...
>
> dat <- data.frame(ftable(tapply(clock.df$has345, list(clock.df$season, clock.df$tra
p, clock.df$sex, clock.df$has345), length)))
> names(dat) <- c("season", "trap", "sex", "has345", "freq")
> dat$freq[is.na(dat$freq) == TRUE] <- 0
> y <- cbind(dat[, 5][dat[, 4] == "Y"], dat[, 5][dat[, 4] == "N"])
> data <- dat[1:18, 1:3]
>
> model.global <- glm(y ~ season + trap + sex + season:sex + trap:sex +
trap:season, data = data, family = quasibinomial)
> summary(model.global)

Call:
glm(formula = y ~ season + trap + sex + season:sex + trap:sex +
    trap:season, family = quasibinomial, data = data)

Deviance Residuals:
    1      2      3      4      5      6      7
-0.39759  0.06773  0.30353 -1.25787  0.51615 -0.10285  0.55963
0.30964
    9     10     11     12     13     14     15
-0.17966  0.49262 -0.05553 -0.39731  0.80819 -0.46759  0.08391
0.56446
    17     18
 0.28161  0.19244

Coefficients:
(Intercept)   -0.31669   Std. Error 0.29068   t value -1.089   Pr(>|t|) 0.3372
season2         0.56001   Std. Error 0.39021   t value  1.435   Pr(>|t|) 0.2246
season3         0.85176   Std. Error 0.37418   t value  2.276   Pr(>|t|) 0.0851
trap2          -0.88148   Std. Error 0.91478   t value -0.964   Pr(>|t|) 0.3898
trap3          -0.62675   Std. Error 0.35727   t value -1.754   Pr(>|t|) 0.1542
```

Appendix

```
sexM          -0.22509    0.37937   -0.593    0.5849
season2:sexM   -0.09245    0.40288   -0.229    0.8298
season3:sexM   -0.84231    0.40922   -2.058    0.1087
trap2:sexM      1.16122    0.46247    2.511    0.0660 .
trap3:sexM      0.49351    0.34752    1.420    0.2286
season2:trap2   0.77363    0.96831    0.799    0.4691
season3:trap2   0.61834    0.93510    0.661    0.5446
season2:trap3   0.53383    0.43005    1.241    0.2823
season3:trap3   1.06392    0.43725    2.433    0.0717 .
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 0.9010412)

Null deviance: 54.5102 on 17 degrees of freedom
Residual deviance: 4.2728 on 4 degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 4

```
>
> model1.1 <- update(model.global, ~.-season:sex)
> summary(model1.1)
```

```
Call:
glm(formula = y ~ season + trap + sex + trap:sex + season:trap,
     family = quasibinomial, data = data)
```

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -1.45097 | -0.64970 | -0.00137 | 0.61464 | 1.03751 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|---------------|----------|------------|---------|----------|
| (Intercept) | -0.1933 | 0.3500 | -0.552 | 0.601 |
| season2 | 0.5761 | 0.4405 | 1.308 | 0.239 |
| season3 | 0.5356 | 0.4345 | 1.233 | 0.264 |
| trap2 | -0.6734 | 1.1417 | -0.590 | 0.577 |
| trap3 | -0.5918 | 0.4543 | -1.303 | 0.240 |
| sexM | -0.5501 | 0.3549 | -1.550 | 0.172 |
| trap2:sexM | 0.8878 | 0.5721 | 1.552 | 0.172 |
| trap3:sexM | 0.4962 | 0.4464 | 1.111 | 0.309 |
| season2:trap2 | 0.7070 | 1.2235 | 0.578 | 0.584 |
| season3:trap2 | 0.4576 | 1.1773 | 0.389 | 0.711 |
| season2:trap3 | 0.4864 | 0.5553 | 0.876 | 0.415 |
| season3:trap3 | 0.9891 | 0.5559 | 1.779 | 0.125 |

(Dispersion parameter for quasibinomial family taken to be 1.490203)

Null deviance: 54.510 on 17 degrees of freedom
Residual deviance: 9.749 on 6 degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 4

```
>
> anova(model.global,model1.1,test="F")
Analysis of Deviance Table
```

```
Model 1: y ~ season + trap + sex + season:sex + trap:sex + trap:season
Model 2: y ~ season + trap + sex + trap:sex + season:trap
```

| | Resid. | Df | Resid. Dev | Df | Deviance | F | Pr(>F) |
|---|--------|----|------------|----|----------|--------|--------|
| 1 | | 4 | 4.2728 | | | | |
| 2 | | 6 | 9.7490 | -2 | -5.4762 | 3.0388 | 0.1575 |

```
>
> model1.2 <- update(model1.1, ~.-season:trap)
```

Appendix

```
> summary(model.2)
```

```
Call:
glm(formula = y ~ season + trap + sex + trap:sex, family = quasibinomial,
    data = data)
```

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|--------|--------|--------|
| -1.7746 | -0.6739 | 0.1022 | 0.5487 | 1.7578 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|------------|
| (Intercept) | -0.51951 | 0.29070 | -1.787 | 0.10422 |
| season2 | 0.93988 | 0.26203 | 3.587 | 0.00495 ** |
| season3 | 1.09733 | 0.26070 | 4.209 | 0.00180 ** |
| trap2 | -0.31689 | 0.42942 | -0.738 | 0.47751 |
| trap3 | -0.05838 | 0.30664 | -0.190 | 0.85281 |
| sexM | -0.56779 | 0.36070 | -1.574 | 0.14653 |
| trap2:sexM | 0.90106 | 0.58057 | 1.552 | 0.15171 |
| trap3:sexM | 0.48919 | 0.44815 | 1.092 | 0.30062 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 1.538733)

Null deviance: 54.510 on 17 degrees of freedom
Residual deviance: 16.016 on 10 degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 4

```
>
> anova(model.1,model.2,test="F")
Analysis of Deviance Table
```

```
Model 1: y ~ season + trap + sex + trap:sex + season:trap
Model 2: y ~ season + trap + sex + trap:sex
  Resid. Df Resid. Dev Df Deviance    F Pr(>F)
1         6      9.749
2        10     16.015  -4   -6.2665 1.0513 0.4545
```

```
>
> model.3 <- update(model.2, ~.-trap:sex)
> summary(model.3)
```

```
Call:
glm(formula = y ~ season + trap + sex, family = quasibinomial,
    data = data)
```

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|--------|--------|--------|
| -2.3822 | -0.7702 | 0.1004 | 0.7615 | 1.4521 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|------------|
| (Intercept) | -0.7024 | 0.2643 | -2.657 | 0.02089 * |
| season2 | 0.9130 | 0.2645 | 3.451 | 0.00479 ** |
| season3 | 1.0941 | 0.2640 | 4.144 | 0.00136 ** |
| trap2 | 0.1605 | 0.2991 | 0.536 | 0.60145 |
| trap3 | 0.1686 | 0.2257 | 0.747 | 0.46936 |
| sexM | -0.1461 | 0.1983 | -0.737 | 0.47551 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 1.588046)

Null deviance: 54.510 on 17 degrees of freedom

Appendix

Residual deviance: 19.948 on 12 degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 4

```
>
> anova(model1.2,model1.3,test="F")
Analysis of Deviance Table

Model 1: y ~ season + trap + sex + trap:sex
Model 2: y ~ season + trap + sex
  Resid. Df Resid. Dev Df Deviance      F Pr(>F)
1         10      16.015
2         12      19.948 -2   -3.9328 1.2779 0.3205
>
> model1.4 <- update(model1.3, ~.-sex)
> summary(model1.4)
```

Call:
glm(formula = y ~ season + trap, family = quasibinomial, data = data)

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|---------|---------|---------|
| -2.60490 | -0.72283 | 0.04899 | 0.95417 | 1.71171 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|-------------|----------|------------|---------|----------|-----|
| (Intercept) | -0.7647 | 0.2464 | -3.104 | 0.00838 | ** |
| season2 | 0.8968 | 0.2586 | 3.467 | 0.00417 | ** |
| season3 | 1.0966 | 0.2591 | 4.232 | 0.00098 | *** |
| trap2 | 0.1382 | 0.2921 | 0.473 | 0.64399 | |
| trap3 | 0.1647 | 0.2214 | 0.744 | 0.47025 | |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 1.531983)

Null deviance: 54.510 on 17 degrees of freedom
Residual deviance: 20.811 on 13 degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 3

```
>
> anova(model1.3,model1.4,test="F")
Analysis of Deviance Table

Model 1: y ~ season + trap + sex
Model 2: y ~ season + trap
  Resid. Df Resid. Dev Df Deviance      F Pr(>F)
1         12      19.948
2         13      20.811 -1   -0.86271 0.5432 0.4753
>
> model1.5 <- update(model1.4, ~.-trap)
> summary(model1.5)
```

Call:
glm(formula = y ~ season, family = quasibinomial, data = data)

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|---------|---------|---------|
| -2.89968 | -0.61665 | 0.04609 | 0.62619 | 1.91790 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|-------------|----------|------------|---------|----------|----|
| (Intercept) | -0.6581 | 0.1911 | -3.444 | 0.003615 | ** |

Appendix

```
season2      0.8999      0.2449      3.674 0.002257 **
season3      1.1043      0.2386      4.628 0.000328 ***
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 1.394941)

```
Null deviance: 54.510 on 17 degrees of freedom
Residual deviance: 21.686 on 15 degrees of freedom
AIC: NA
```

Number of Fisher Scoring iterations: 3

```
>
> anova(model.4,model.5,test="F")
Analysis of Deviance Table
```

```
Model 1: y ~ season + trap
Model 2: y ~ season
      Resid. Df Resid. Dev Df Deviance      F Pr(>F)
1          13      20.811
2          15      21.686 -2  -0.87455  0.2854 0.7563
```

```
>
>
> model.6 <- update(model.4, ~.-season)
> summary(model.6)
```

```
Call:
glm(formula = y ~ trap, family = quasibinomial, data = data)
Deviance Residuals:
```

```
      Min       1Q   Median       3Q      Max
-3.2547  -1.0711  -0.0279   0.8884   3.3154
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.04832    0.25188  -0.192   0.850
trap2        0.40876    0.41052   0.996   0.335
trap3        0.15893    0.31593   0.503   0.622
```

(Dispersion parameter for quasibinomial family taken to be 3.281313)

```
Null deviance: 54.510 on 17 degrees of freedom
Residual deviance: 51.231 on 15 degrees of freedom
AIC: NA
```

Number of Fisher Scoring iterations: 3

```
> anova(model.4,model.6, test = "F")
Analysis of Deviance Table
```

```
Model 1: y ~ season + trap
Model 2: y ~ trap
      Resid. Df Resid. Dev Df Deviance      F    Pr(>F)
1          13      20.811
2          15      51.231 -2   -30.42  9.9282 0.002413 **
```

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> s1<-coef(model.5)[[1]]
> s2<-coef(model.5)[[1]] + coef(model.5)[[2]]
> s3<-coef(model.5)[[1]] + coef(model.5)[[3]]
> s <- c(s1,s2,s3)
> 1/(1+1/(exp(s))) # after back transforming from logit: gives mean
proportion estimated by the GLM for 345 being present in each season;
season1: 0.3411765; season2: 0.5601660; season3: 0.6097561
[1] 0.3411765 0.5601660 0.6097561
>
```

Appendix

```
> gof<-cbind(res.diviance=model.5$deviance, df=model.5$df.residual, p=1-  
pchisq(model.5$deviance, model.5$df.residual))  
> gof  
      res.diviance df          p  
[1,]      21.68561 15 0.1163164
```


References

- Abadia-Cardoso A, Clemento AJ, Garza JC (2011) Discovery and characterization of single-nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. *Molecular Ecology Resources*, **11**, 31-49.
- Aguilar A, Garza JC (2008) Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). *Molecular Ecology Resources*, **8**, 659-662.
- Allendorf FW (1986) Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, **5**, 181-190.
- Allendorf FW, England PR, Luikart G, Ritchie PA, Ryman N (2008) Genetic effects of harvest on wild animal populations. *Trends in Ecology & Evolution*, **23**.
- Allendorf FW, Luikart G (2007) *Conservation and The Genetics of Populations*. Oxford, Blackwell Publishing Ltd.
- Allendorf FW, Lundquist LL (2003) Introduction: Population Biology, Evolution, and Control of Invasive Species. *Conservation Biology*, **17**, 24-30.
- Allendorf FW, Thorgaard GH (1984) Tetraploidy and the evolution of salmonid fishes. 1-53 in T. B. *Evolutionary genetics of fishes*. Plenum Press, New York.
- Allendorf FW, Waples RS (1996) Conservation and Genetics of Salmonid Fishes. 238-280 in J. C. Avise and J. L. Hamrick. *Conservation genetics: Case histories from nature*. Chapman and Hall, New York.
- Anderson EC, Slatkin M (2007) Estimation of the number of individuals founding colonized populations. *Evolution*, **61**, 972-983.
- Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, **9**, 323.
- Arnesen AM, Toften H, Agustsson T, Stefansson SO, Handeland SO, Björnsson BT (2003) Osmoregulation, feed intake, growth and growth hormone levels in 0+ Atlantic salmon (*Salmo salar* L.) transferred to seawater at different stages of smolt development. *Aquaculture*, **222**, 167-187.
- Astaneï I, Gosling E, Wilson J, Powell E (2005) Genetic variability and phylogeography of the invasive zebra mussel, *Dreissena polymorpha* (Pallas). *Molecular Ecology*, **14**, 1655-1666.
- Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155-165.

References

- Banks MA, Blouin MS, Baldwin BA, Rashbrook VK, Fitzgerald HA, Blankenship SM, Hedgecock D (1999) Isolation and inheritance of novel microsatellites in Chinook salmon (*Oncorhynchus tshawytscha*). *Journal of Heredity*, **90**, 281-288.
- Beacham TD (1990) A genetic-analysis of meristic and morphometric variation in chum salmon (*Onchorynchus keta*) at 3 different temperatures. *Canadian Journal of Zoology - Revue Canadienne de Zoologie*, **68**, 225-229.
- Becker LA, Pascual MA, Basso NG (2007) Colonization of the southern patagonia ocean by exotic Chinook salmon. *Conservation Biology*, **21**, 1347-1352.
- Behnke RJ (1992) *Native Trout of Western North America*. Bethesda, Maryland, American Fisheries Society.
- Behnke RJ (2002) Comment: First documented case of anadromy in a population of introduced rainbow trout in Patagonia, Argentina. *Transactions of the American Fisheries Society*, **131**, 582-585.
- Bell-Pedersen D, Cassone VM, Earnest DJ, Golden SS, Hardin PE, Thomas TL, Zoran MJ (2005) Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics*, **6**, 544-556.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, 289-300.
- Berryman AA (2002) Population: a central concept for ecology? *Oikos*, **97**, 439-442.
- Bromage N, Porter M, Randall C (2001) The environmental regulation of maturation in farmed finfish with special reference to the role of photoperiod and melatonin. *Aquaculture*, **197**, 63-98.
- Brown JE, Stepien CA (2009) Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. *Molecular Ecology*, **18**, 64-79.
- Brown JW, Van Coeverden De Groot PJ, Birt TP, Seutin G, Boag PT, Friesen VL (2007) Appraisal of the consequences of the DDT-induced bottleneck on the level and geographic distribution of neutral genetic variation in Canadian peregrine falcons, *Falco peregrinus*. *Molecular Ecology*, **16**, 327-343.
- Buchholz WG, Miller SJ, Spearman WJ (2001) Isolation and characterization of chum salmon microsatellite loci and use across species. *Animal Genetics*, **32**, 162-165.
- Busack CA, Gall GAE (1980) Ancestry of artificially propagated California rainbow trout strains. *California Fish and Game*, **66**, 17-24.

References

- Cairney M, Taggart JB, Hoyheim B (2000) Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. *Molecular Ecology*, **9**, 2175-2178.
- Campbell NR, Overturf KEN, Narum SR (2009) Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. *Molecular Ecology Resources*, **9**, 318-322.
- Carlson D, Williams D (1999) Sex-specific vitellogenin production in immature rainbow trout. *Environ Toxicol Chem*, **18**, 2361 - 2363.
- Carlson SM, Seamons TR (2008) A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. *Evolutionary Applications*, **1**, 222-238.
- Carvalho GR, Hauser L (1994) Molecular genetics and the stock concept in fisheries. *Reviews in Fish Biology and Fisheries*, **4**, 326-350.
- Caudron A, Champigneulle A, Guyomard R (2009) Evidence of two contrasting brown trout *Salmo trutta* populations spatially separated in the River Borne (France) and shift in management towards conservation of the native lineage. *Journal of Fish Biology*, **74**, 1070-1085.
- Chakraborty R, Ryman N (1983) Relationship of Mean and Variance of Genotypic Values with Heterozygosity Per Individual in a Natural-Population. *Genetics*, **103**, 149-152.
- Charlesworth B, Charlesworth D, Barton NH (2003) The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 99-125.
- Clemento AJ, Abadía-Cardoso A, Starks HA, Garza JC (2011) Discovery and characterization of single nucleotide polymorphisms in Chinook salmon, *Oncorhynchus tshawytscha*. *Molecular Ecology Resources*, **11**, 50-66.
- Colautti RI, Manca M, Viljanen M, Ketelaars HAM, Bürgi H, Macisaac HJ, Heath DD (2005) Invasion genetics of the Eurasian spiny waterflea: evidence for bottlenecks and gene flow using microsatellites. *Molecular Ecology*, **14**, 1869-1879.
- Conover DO (1998) Local adaptation in marine fishes: evidence and implications for stock enhancement. *Bulletin of Marine Science*, **62**, 477-493.
- Conover DO, Munch SB (2002) Sustaining fisheries yields over evolutionary time scales. *Science*, **297**, 94-96.
- Consuegra S, Garcia de Leaniz C, Serdio A, Verspoor E (2005) Selective exploitation of early running fish may induce genetic and phenotypic changes in Atlantic salmon. *Journal of Fish Biology*, **67**, 129-145.

References

- Copp GH, Bianco PG, Bogutskaya NG, Erős T, Falka I, Ferreira MT, Fox MG, Freyhof J, Gozlan RE, Grabowska J, Kováč V, Moreno-Amich R, Naseka AM, Peñáz M, Povž M, Przybylski M, Robillard M, Russell IC, Stakėnas S, Šumer S, Vila-Gispert A, Wiesner C (2005a) To be, or not to be, a non-native freshwater fish? *Journal of Applied Ichthyology*, **21**, 242-262.
- Copp GH, Garthwaite R, Gozlan RE (2005b) Risk identification and assessment of non-native freshwater fishes: a summary of concepts and perspectives on protocols for the UK. *Journal of Applied Ichthyology*, **21**, 371-373.
- Crowl TA, Townsend CR, McIntosh AR (1992) The impact of introduced brown and rainbow trout on native fish: the case of Australasia. *Reviews in Fish Biology and Fisheries*, **2**, 217-241.
- Cuveliers EL, Volckaert FAM, Rijnsdorp AD, Larmuseau MHD, Maes GE (2011) Temporal genetic stability and high effective population size despite fisheries-induced life-history trait evolution in the North Sea sole. *Molecular Ecology*, **20**, 3555-3568.
- Danzmann RG, Ferguson MM, Heculuck DM (1994) Heterogeneity in the distribution of mitochondrial DNA haplotypes in female rainbow trout spawning in different seasons. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 284-289.
- Darimont CT, Carlson SM, Kinnison MT, Paquet PC, Reimchen TE, Wilmers CC (2009) Human predators outpace other agents of trait change in the wild. *Proceedings of the National Academy of Sciences*, **106**, 952-954.
- Darlington TK, Wager-Smith K, Ceriani MF, Staknis D, Gekakis N, Steeves TDL, Weitz CJ, Takahashi JS, Kay SA (1998) Closing the circadian loop: *CLOCK*-Induced Transcription of its own inhibitors *per* and *tim*. *Science*, **280**, 1599-1603.
- Davie A, Minghetti M, Migaud H (2009) Seasonal Variations in Clock- Gene Expression in Atlantic Salmon (*Salmo salar*). *Chronobiology International*, **26**, 379-395.
- Dedual M (2002) Lipid content in rainbow trout (*Oncorhynchus mykiss*) fry and parr reared in spawning tributaries of Lake Taupo, New Zealand. *New Zealand journal of marine and freshwater research*, **36**, 809-814.
- DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 461-473.
- Docker MF, Heath DD (2003) Genetic comparison between sympatric anadromous steelhead and freshwater resident rainbow trout in British Columbia, Canada. *Conservation Genetics*, **4**, 227-231.
- Dollar AM, Katz M (1964) Rainbow trout brood stocks and strains in American hatcheries as factors in the occurrence of hepatoma. *The Progressive Fish-Culturist*, **1964**, 167-174.

References

- Duncan JR, Lockwood JL (2001) Extinction in a field of bullets: a search for causes in the decline of the world's freshwater fishes. *Biological Conservation*, **102**, 97-105.
- Dunlop ES, Shuter BJ, Dieckmann U (2007) Demographic and evolutionary consequences of selective mortality: Predictions from an eco-genetic model for smallmouth bass. *Transactions of the American Fisheries Society*, **136**, 749-765.
- Duston J, Bromage N (1988) The entrainment and gating of the endogenous circannual rhythm of reproduction in the female rainbow trout (*Salmo gairdneri*). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, **164**, 259-268.
- Dwyer WP, Piper RG (1984) Three-year hatchery and field evaluation of four strains of rainbow trout. *North American Journal of Fisheries Management*, **4**, 216-221.
- Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences*, **97**, 7043-7050.
- England PR, Osler GHR, Woodworth LM, Montgomery ME, Briscoe DA, Frankham R (2003) Effects of intense versus diffuse population bottlenecks on microsatellite genetic diversity and evolutionary potential. *Conservation Genetics*, **4**, 595-604.
- Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113-4130.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611-2620.
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.
- Foll M, Gaggiotti OE (2008) Identifying the environmental factors that determine the genetic structure of populations. *Genetics*, **174**, 875-891.
- Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131-140.
- Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity*, **106**, 404-420.
- Gall GAE, Crandell PA (1992) The rainbow trout. *Aquaculture*, **100**, 1-10.
- Gall GAE, Huang N (1988) Heritability and selection schemes for rainbow trout: body weight. *Aquaculture*, **73**, 43-56.

References

- Gekakis N, Staknis D, Nguyen H, Davis F, Wilsbacher L, King D, Takahashi J, Weitz C (1998) Role of the *CLOCK* protein in the mammalian circadian mechanism. *Science*, **280**, 1564 - 1569.
- Glover K, Hansen M, Lien S, Als T, Hoyheim B, Skaala O (2010) A comparison of SNP and STR loci for delineating population structure and performing individual genetic assignment. *BMC Genetics*, **11**, 2.
- Goudet J (1995) Fstat version 1.2: a computer program to calculate F statistics. *Journal of Heredity*, **86**, 485-486.
- Greig C, Banks MA (1999) Five multiplexed microsatellite loci for rapid response run identification of California's endangered winter chinook salmon. *Animal Genetics*, **30**, 318-320.
- Groot C, Margolis L (1991) *Pacific salmon life histories*, University of British Columbia Press.
- Gyllensten U (1985) The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous, and freshwater species. *Journal of Fish Biology*, **26**, 691-699.
- Halverson A (2010) *An Entirely Synthetic Fish*. New Haven, Yale University Press.
- Hänfling B (2007) Understanding the establishment success of non-indigenous fishes: lessons from population genetics. *Journal of Fish Biology*, **71**, 115-135.
- Hansen MM (2002) Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: an approach using microsatellite DNA analysis of historical and contemporary samples. *Molecular Ecology*, **11**, 1003-1015.
- Hard JJ, Gross MR, Heino M, Hilborn R, Kope RG, Law R, Reynolds JD (2008) Evolutionary consequences of fishing and their implications for salmon. *Evolutionary Applications*, **1**, 388-408.
- Hartl DL, Clark AG (1997) *Principles of Population Genetics*. Sunderland, Massachusetts, Sinauer Associates, Inc.
- Hauser L, Adcock GJ, Smith PJ, Bernal Ramírez JH, Carvalho GR (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences*, **99**, 11742-11747.
- Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, **9**, 333-362.

References

- Heath DD, Busch C, Kelly J, Atagi DY (2002) Temporal change in genetic structure and effective population size in steelhead trout (*Oncorhynchus mykiss*). *Molecular Ecology*, **11**, 197-214.
- Heath DD, Pollard S, Herbinger C (2001) Genetic structure and relationships among steelhead trout (*Oncorhynchus mykiss*) populations in British Columbia. *Heredity*, **86**, 618-627.
- Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313-318.
- Heggberget TG (1988) Timing of spawning in Norwegian Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **45**, 845-849.
- Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, **290**, 516-518.
- Hess JE, Narum SR (2011) Single-nucleotide polymorphism (SNP) loci correlated with run timing in adult Chinook salmon from the Columbia River Basin. *Transactions of the American Fisheries Society*, **140**, 855-864.
- Hobbs DF (1948) *Trout fisheries in New Zealand: their development and management*. Wellington, N.Z., N.Z. Marine Dept.
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65-70.
- Huey RB, Gilchrist GW, Carlson ML, Berrigan D, Serra L (2000) Rapid evolution of a geographic cline in size in an introduced fly. *Science*, **287**, 308-309.
- Hurlbert SH, Zedler J, Fairbanks D (1972) Ecosystem alteration by mosquitofish (*Gambusia affinis*) predation. *Science*, **175**, 639-641.
- Ihaka R, Gentleman R (1996) R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, **5**, 314.
- Johnsen A, Fidler AE, Kuhn S, Carter KL, Hoffmann A, Barr IR, Biard C, Charmantier A, Eens M, Korsten P, Siitari H, Tomiuk J, Kempenaers B (2007) Avian Clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Molecular Ecology*, **16**, 4867-4880.
- Jorgensen C, Enberg K, Dunlop ES, Arlinghaus R, Boukal DS, Brander K, Ernande B, Gårdmark A, Johnston F, Matsumura S, Pardoe H, Raab K, Silva A, Vainikka A, Dieckmann U, Heino M, Rijnsdorp AD (2007) Managing evolving fish stocks. *Science*, **318**, 1247-1248.

References

- Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes*, **5**, 187-189.
- Kalinowski ST, Muhlfeld CC, Guy CS, Cox B (2010) Founding population size of an aquatic invasive species. *Conservation Genetics*, **11**, 2049-2053.
- Kawamura K, Yonekura R, Katano O, Taniguchi Y, Saitoh K (2006) Origin and dispersal of bluegill sunfish, *Lepomis macrochirus*, in Japan and Korea. *Molecular Ecology*, **15**, 613-621.
- Kinnison MT, Bentzen P, Unwin MJ, Quinn TP (2002) Reconstructing recent divergence: evaluating nonequilibrium population structure in New Zealand Chinook salmon. *Molecular Ecology*, **11**, 739-754.
- Kinnison MT, Quinn TP, Unwin MJ (2011) Correlated contemporary evolution of life history traits in New Zealand Chinook salmon, *Oncorhynchus tshawytscha*. *Heredity*, **106**, 448-459.
- Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution*, **16**, 199-204.
- Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177-181.
- Krueger CC, May B (1987) Genetic comparison of naturalized rainbow trout populations among Lake Superior tributaries: Differentiation based on allozyme data. *Transactions of the American Fisheries Society*, **116**, 795-806.
- Kyriacou CP, Peixoto AA, Sandrelli F, Costa R, Tauber E (2008) Clines in clock genes: fine-tuning circadian rhythms to the environment. *Trends in Genetics*, **24**, 124-132.
- Laikre L, Palm S, Ryman N (2005) Genetic population structure of fishes: Implications for coastal zone management. *Ambio*, **34**, 111-119.
- Lamberti GA, Gregory SV, Ashkenas LR, Wildman RC, Moore KMS (1991) Stream ecosystem recovery following a catastrophic debris flow. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 196-208.
- Leary RF, Allendorf FW, Knudsen KL (1989) Genetic differences among rainbow trout spawned on different days within a single season. *The Progressive Fish-Culturist*, **51**, 10-19.
- Leder EH, Danzmann RG, Ferguson MM (2006) The candidate gene, *Clock*, localizes to a strong spawning time quantitative trait locus region in rainbow trout. *Journal of Heredity*, **97**, 74-80.
- Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, **17**, 386-391.

References

- Liedvogel M, Szulkin M, Knowles SCL, Wood MJ, Sheldon BC (2009) Phenotypic correlates of *Clock* gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction. *Molecular Ecology*, **18**, 2444-2456.
- Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998a) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238-247.
- Luikart G, Cornuet J-M (1998) Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*, **12**, 228-237.
- Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998b) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**, 963-974.
- MacCrimmon HR (1971) World distribution of rainbow trout (*Salmo gairdneri*). *Journal of the Fisheries Research Board of Canada*, **28**, 663-704.
- Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, **25**, 1982-1983.
- McClelland EK, Naish KA (2007) Comparisons of FST and QST of growth-related traits in two populations of Coho salmon. *Transactions of the American Fisheries Society*, **136**, 1276-1284.
- McConnell S, Hamilton L, Morris D, Cook D, Paquet D, Bentzen P, Wright J (1995) Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. *Aquaculture*, **137**, 19-30.
- McCusker MR, Bentzen P (2010) Positive relationships between genetic diversity and abundance in fishes. *Molecular Ecology*, **19**, 4852-4862.
- McDonald JH (1994) Detecting natural selection by comparing geographic variation in protein and DNA polymorphisms. 88–100 in B. Golding. *Non-neutral evolution: theories and molecular data*. Chapman and Hall, New York.
- McDowall RM (1991) Freshwater fisheries research in New Zealand: Processes, projects, and people. *New Zealand Journal of Marine and Freshwater Research*, **25**, 393-413.
- McDowall RM (1994) *Gamekeepers for the Nation*. Christchurch, Canterbury University Press.
- McDowall RM (2005) Letter to the editor. *New Zealand Journal of Marine and Freshwater Research*, **39**, 1199-1200.

References

- McPhee MV, Quinn TP (1998) Factors affecting the duration of nest defense and reproductive lifespan of female sockeye salmon, *Oncorhynchus nerka*. *Environmental Biology of Fishes*, **51**, 369-375.
- Miller LM, Kapuscinski AR (1997) Historical Analysis of Genetic Variation Reveals Low Effective Population Size in a Northern Pike (*Esox lucius*) Population. *Genetics*, **147**, 1249-1258.
- Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences*, **98**, 5446-5451.
- Morris DB, Richard KR, Wright JM (1996) Microsatellites from rainbow trout (*Oncorhynchus mykiss*) and their use for genetic study of salmonids. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 120-126.
- Mousseau TA, Roff DA (1987) Natural selection and the heritability of fitness components. *Heredity*, **59**, 181-197.
- Nakagawa S (2004) A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*, **15**, 1044-1045.
- Narum S (2006) Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, **7**, 783-787.
- Narum SR, Banks MA, Beacham TD, Bellinger MR, Campbell MR, Dekoning J, Elz A, Guthrie III CM, Kozfkay C, Miller KM, Moran P, Phillips R, Seeb LW, Smith CT, Warheit K, Young SF, Garza JC (2008) Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. *Molecular Ecology*, **17**, 3464-3477.
- Nelson RJ, Beacham TD (1999) Isolation and cross species amplification of microsatellite loci useful for study of Pacific salmon. *Animal Genetics*, **30**, 228-229.
- Nielsen EE, Hansen MM, Loeschcke V (1997) Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: A comparison of genetic composition over 60 years. *Molecular Ecology*, **6**, 487-492.
- Nielsen EE, Hansen MM, Loeschcke V (1999a) Analysis of DNA from old scale samples: Technical aspects, applications and perspectives for conservation. *Heredity*, **130**, 265-276.
- Nielsen EE, Hansen MM, Loeschcke V (1999b) Genetic variation in time and space: Microsatellite analysis of extinct and extant populations of Atlantic salmon. *Evolution*, **53**, 261-268.
- Northwest Regional Office N. (2011) "Steelhead Distinct Population Segments and their ESA Listing Status." Retrieved 4/09/2011, 2011, from <http://www.nwr.noaa.gov/ESA-Salmon-Listings/Salmon-Populations/Steelhead/>.

References

- O'Connell M, Danzmann RG, Cornuet J-M, Wright JM, Ferguson MM (1997) Differentiation of rainbow trout (*Oncorhynchus mykiss*) populations in Lake Ontario and the evaluation of the stepwise mutation and infinite allele mutation models using microsatellite variability. *Canadian Journal of Fisheries and Aquatic Sciences*, **54**, 1391-1399.
- O'Malley K, Banks M (2008a) A latitudinal cline in the Chinook salmon (*Oncorhynchus tshawytscha*) Clock gene: evidence for selection on PolyQ length variants. *Proceedings of the Royal Society B*, **275**, 2813 - 2821.
- O'Malley K, Camara M, Banks M (2007) Candidate loci reveal genetic differentiation between temporally divergent migratory runs of Chinook salmon (*Oncorhynchus tshawytscha*). *Molecular Ecology*, **16**, 4930 - 4941.
- O'Malley KG, Banks MA (2008b) Duplicated *Clock* genes with unique polyglutamine domains provide evidence for nonhomologous recombination in Chinook salmon (*Oncorhynchus tshawytscha*). *Genetica*, **132**, 87-94.
- O'Malley KG, Ford MJ, Hard JJ (2010) *Clock* polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 3703-3714.
- Olsen J, Wuttig K, Fleming D, Kretschmer E, Wenburg J (2006) Evidence of partial anadromy and resident-form dispersal bias on a fine scale in populations of *Oncorhynchus mykiss*. *Conservation Genetics*, **7**, 613-619.
- Olsen JB, Bentzen P, Banks MA, Shaklee JB, Young S (2000) Microsatellites reveal population identity of individual pink salmon to allow supportive breeding of a population at risk of extinction. *Transactions of the American Fisheries Society*, **129**, 232-242.
- Olsen JB, Bentzen P, Seeb JS (1998) Characterization of seven microsatellite loci derived from pink salmon. *Molecular Ecology*, **7**, 1087-1089.
- Paibomesai M, Moghadam H, Ferguson M, Danzmann R (2010) *Clock* genes and their genomic distributions in three species of salmonid fishes: Associations with genes regulating sexual maturation and cell cycling. *BMC Research Notes*, **3**, 215.
- Palsbøll PJ, Bérubé M, Allendorf FW (2007) Identification of management units using population genetic data. *Trends in Ecology & Evolution*, **22**, 11-16.
- Pascual MA, Bentzen P, Rossi CR, Mackey G, Kinnison MT, Walker R (2001) First documented case of anadromy in a population of introduced rainbow trout in Patagonia, Argentina. *Transactions of the American Fisheries Society*, **130**, 53-67.
- Peakall ROD, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

References

- Pearse D, Martinez E, Garza J (2011) Disruption of historical patterns of isolation by distance in coastal steelhead. *Conservation Genetics*, **12**, 691-700.
- Pearse DE, Garza JC (2008). Historical baseline for genetic monitoring of coastal California steelhead, *Oncorhynchus mykiss*. California Department of Fish and Game Fisheries Restoration Grant Program
- Pertoldi C, Hansen MM, Loeschcke V, Madsen AB, Jacobsen L, Baagoe H (2001) Genetic consequences of population decline in the European otter (*Lutra lutra*): an assessment of microsatellite DNA variation in Danish otters from 1883 to 1993. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 1775-1781.
- Piry S, Luikart G, Cornuet JM (1999) *Bottleneck : A Computer Program for Detecting Recent Reductions in the Effective Population Size Using Allele Frequency Data*. Cary, NC, ETATS-UNIS, Oxford University Press.
- Pringle RM (2005) The origins of the Nile perch in Lake Victoria. *BioScience*, **55**, 780-787.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
- Queney G, Ferrand N, Marchandeu S, Azevedo M, Mougel F, Branco M, Monnerot M (2000) Absence of a genetic bottleneck in a wild rabbit (*Oryctolagus cuniculus*) population exposed to a severe viral epizootic. *Molecular Ecology*, **9**, 1253-1264.
- Quinn TP (2005) *The Behavior and Ecology of Pacific Salmon and Trout*. Bethesda, Maryland, American Fisheries Society.
- Quinn TP, Adams DJ (1996) Environmental changes affecting the migratory timing of American shad and Sockeye salmon. *Ecology*, **77**, 1151-1162.
- Quinn TP, Hodgson S, Flynn L, Hilborn R, Rogers DE (2007) Directional selection by fisheries and the timing of sockeye salmon (*Oncorhynchus nerka*) migrations. *Ecological Applications*, **17**, 731-739.
- Quinn TP, McGinnity P, Cross TF (2006) Long-term declines in body size and shifts in run timing of Atlantic salmon in Ireland. *Journal of Fish Biology*, **68**, 1713-1730.
- Quinn TP, Nielsen JL, Gan C, Unwin MJ, Wilmot R, Guthrie C, Utter FM (1996) Origin and genetic structure of Chinook salmon, *Oncorhynchus tshawytscha*, transplanted from California to New Zealand: allozyme and mtDNA evidence. *Fishery Bulletin*, **94**, 506-521.
- Quinn TP, Peterson JA, Galluci VF, Hershberger WK, Brannon EL (2002) Artificial selection and environmental change: Countervailing factors affecting the timing of spawning by coho and chinook salmon. *Transactions of the American Fisheries Society*, **131**, 591-598.

References

- Quinn TP, Unwin MJ, Kinnison MT (2000) Evolution of Temporal Isolation in the Wild: Genetic Divergence in Timing of Migration and Breeding by Introduced Chinook Salmon Populations. *Evolution*, **54**, 1372-1385.
- Raymond M, Rousset F (1995) Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity*, **86**, 248-249.
- Rexroad CE, Coleman RL, Gustafson AL, Hershberger WK, Killefer J (2002) Development of rainbow trout microsatellite markers from repeat enriched libraries. *Marine Biotechnology*, **4**, 12-16.
- Rexroad CE, Palti Y, Gahr SA, Vallejo RL (2008) A second generation genetic map for rainbow trout (*Oncorhynchus mykiss*). *BMC Genet*, **9**, 74.
- Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica*, **112-113**, 183-198.
- Rioux Paquette S (2011). PopGenKit: useful function for (batch) file conversion and data resampling in microsatellite datasets. . [R package](#).
- Riva Rossi CM, Lessa EP, Pascual MA (2004) The origin of introduced rainbow trout (*Oncorhynchus mykiss*) in the Santa Cruz River, Patagonia, Argentina, as inferred from mitochondrial DNA. *Canadian Journal of Fisheries and Aquatic Sciences*, **61**, 1095-1101.
- Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, **22**, 454-464.
- Rosenau ML (1991) Natal-stream rearing in three populations of rainbow trout in Lake Taupo, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, **25**, 81-91.
- Ruesink JL (2005) Global Analysis of Factors Affecting the Outcome of Freshwater Fish Introductions. *Conservation Biology*, **19**, 1883-1893.
- Ruzzante DE, Taggart CT, Doyle RW, Cook D (2001) Stability in the historical pattern of genetic structure of Newfoundland cod (*Gadus morhua*) despite the catastrophic decline in population size from 1964 to 1994. *Conservation Genetics*, **2**, 257-269.
- Sakamoto T (1996). Study for the introduction of positional cloning method using DNA markers in fisheries science. Department of Aquatic Biosciences. Tokyo, Tokyo University of Fisheries. **PhD**.
- Sakamoto T, Danzmann R, Okamoto N, Ferguson M, Ihssen P (1999) Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*, **173**, 33 - 43.

References

- Sanchez JA, Clabby C, Ramos D, Blanco G, Flavin F, Vazquez E, Powell R (1996) Protein and microsatellite single locus variability in Atlantic salmon (*Salmo salar* L). *Heredity*, **77**, 423-432.
- Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution*, **22**, 25-33.
- Scott D, Hewitson J, Fraser JC (1978) The origins of rainbow trout, *Salmo Gairdneri* Richardson, in New Zealand. *California Fish and Game*, **64**, 210-218.
- Scribner KT, Gust JR, Fields RL (1996) Isolation and characterization of novel salmon microsatellite loci: Cross-species amplification and population genetic applications. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 833-841.
- Scrimgeour F, Oxley L (2001) Economic modelling for trout management: an introduction and case study. *Environmental Modelling & Software*, **16**, 571-581.
- Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources*, **11**, 1-8.
- Siitonen L, Gall GAE (1989) Response to selection for early spawn date in rainbow trout, *Salmo gairdneri*. *Aquaculture*, **78**, 153-161.
- Slatkin M (1985) Rare alleles as indicators of gene flow. *Evolution*, **39**, 53-65.
- Smith CT, Koop BF, Nelson RJ (1998) Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. *Molecular Ecology*, **7**, 1614-1616.
- Smith CT, Templin WD, Seeb JE, Seeb LW (2005) Single nucleotide polymorphisms provide rapid and accurate estimates of the proportions of U.S. and Canadian Chinook salmon caught in Yukon River fisheries. *North American Journal of Fisheries Management*, **25**, 944-953.
- Smith GR, Stearley RF (1989) The classification and scientific names of rainbow and cutthroat trouts. *Fisheries*, **14**, 4-10.
- Snowdon RJ, Adam KD (1992) Allozyme variation in rainbow trout from central North Island, New Zealand. *Transactions of the American Fisheries Society*, **121**, 810-814.
- Spies IB, Brasier DJ, O'Reilly TL, Seamons TR, Bentzen P (2005) Development and characterization of novel tetra-, tri-, and dinucleotide microsatellite markers in rainbow trout (*Oncorhynchus mykiss*). *Molecular Ecology Notes*, **5**, 278-281.
- Stephenson J, Campbell M, Hess J, Kozfkay C, Matala A, McPhee M, Moran P, Narum S, Paquin M, Schlei O, Small M, Van Doornik D, Wenburg J (2009) A centralized

References

- model for creating shared, standardized, microsatellite data that simplifies inter-laboratory collaboration. *Conservation Genetics*, **10**, 1145-1149.
- Stephenson RL (1999) Stock complexity in fisheries management: a perspective of emerging issues related to population sub-units. *Fisheries Research*, **43**, 247-249.
- Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, **18**, 94-101.
- Stockwell CA, Mulvey M, Vinyard GL (1996) Translocations and the preservation of allelic diversity. *Conservation Biology*, **10**, 1133-1141.
- Stokell G (1955) *Freshwater fishes of New Zealand*. Christchurch, Simpson and Williams.
- Swanson FJ, Johnson SL, Gregory SV, Acker SA (1998) Flood disturbance in a forested mountain landscape - Interactions of land use and floods. *BioScience*, **48**, 681-689.
- Tallmon DA, Luikart G, Beaumont MA (2004) Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics*, **167**, 977-988.
- Taylor EB (1991) A review of local adaptation in Salmonidac, with particular reference to Pacific and Atlantic salmon. *Aquaculture*, **98**, 185-207.
- Taylor EB, Tamkee P, Sterling G, Hughson W (2007) Microsatellite DNA analysis of rainbow trout (*Oncorhynchus mykiss*) from western Alberta, Canada: native status and evolutionary distinctiveness of 'Athabasca' rainbow trout. *Conservation Genetics*, **8**, 1-15.
- Team RDC (2011) R: A Language and Environment for Statistical Computing.
- Tessier N, Bernatchez L (1999) Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (*Salmo salar*). *Molecular Ecology*, **8**, 169-179.
- Theriault V, Dunlop ES, Dieckmann U, Bernatchez L, Dodson J (2008) The impact of fishing-induced mortality on the evolution of alternative life-history tactics in brook charr. *Evolutionary Applications*, **1**, 409-423.
- Tournier BB, Dardente H, Simonneaux V, Vivien-Roels B, Pévet P, Masson-Pévet M, Vuillez P (2007) Seasonal variations of *Clock* gene expression in the suprachiasmatic nuclei and pars tuberalis of the European hamster (*Cricetus cricetus*). *European Journal of Neuroscience*, **25**, 1529-1536.
- Trépanier S, Rodríguez MA, Magnan P (1996) Spawning migrations in landlocked Atlantic salmon: time series modelling of river discharge and water temperature effects. *Journal of Fish Biology*, **48**, 925-936.

References

- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535-538.
- Venman MR, Dedual M (2005) Migratory behaviour of spawning rainbow trout (*Oncorhynchus mykiss*) in the Tongariro River, New Zealand, after habitat alteration. *New Zealand Journal of Marine and Freshwater Research*, **39**, 951-961.
- Vitaterna MH, Ko CH, Chang A-M, Buhr ED, Fruechte EM, Schook A, Antoch MP, Turek FW, Takahashi JS (2006) The mouse Clock mutation reduces circadian pacemaker amplitude and enhances efficacy of resetting stimuli and phase-response curve amplitude. *Proceedings of the National Academy of Sciences*, **103**, 9327-9332.
- Vitule JRSV, Freire CA, Simberloff D (2009) Introduction of non-native freshwater fish can certainly be bad. *Fish and Fisheries*, **10**, 98-108.
- Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology & Evolution*, **22**, 634-642.
- Wang H (2008) Comparative analysis of teleost fish genomes reveals preservation of different ancient clock duplicates in different fishes. *Marine Genomics*, **1**, 69-78.
- Waples RS (1989) A Generalized Approach for Estimating Effective Population Size From Temporal Changes in Allele Frequency. *Genetics*, **121**, 379-391.
- Waples RS (1991) Pacific salmon, *Oncorhynchus* spp., and the definition of “species” under the Endangered Species Act. *Marine Fisheries Review*, **53**, 11-22.
- Waples RS (1995) Evolutionarily significant units and the conservation of biological diversity under the Endangered Species Act. 8-27 8-27 in J. L. Nielsen and G. A. Powers. *Evolution and the Aquatic Ecosystem: Defining Unique Units in Population Conservation*. **Symposium 17** American Fisheries Society, Bethesda, Maryland.
- Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438-450.
- Waples RS, Do CHI (2008) Ldne: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, **8**, 753-756.
- Waples RS, Gaggiotti O (2006) INVITED REVIEW: 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.
- Waples RS, Gustafson RG, Weitkamp LA, Myers JM, Johnson OW, Busby PJ, Hard JJ, Bryant GJ, Waknitz FW, Nelly K, Teel D, Grant WS, Winans GA, Phelps S, Marshall A, Baker BM (2001) Characterizing diversity in salmon from the Pacific Northwest*. *Journal of Fish Biology*, **59**, 1-41.

References

- Waples RS, Pess GR, Beechie TJ (2008) Evolutionary history of Pacific salmon in dynamic environments. *Evolutionary Applications*, **1**, 189-206.
- Waples RS, Teel DJ (1990) Conservation Genetics of Pacific Salmon .1. Temporal Changes in Allele Frequency. *Conservation Biology*, **4**, 144-156.
- Waples RS, Teel DJ, Myers JM, Marshall AR (2004) Life-history divergence in Chinook salmon: historic contingency and parallel evolution. *Evolution*, **58**, 386-403.
- Ward RD, Woodward M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of Fish Biology*, **44**, 213-232.
- Williamson KS, Cordes JF, May B (2002) Characterization of microsatellite loci in Chinook salmon (*Oncorhynchus tshawytscha*) and cross-species amplification in other salmonids. *Molecular Ecology Notes*, **2**, 17-19.
- Wofford JEB, Gresswell RE, Banks MA (2005) Influence of barriers to movement on within-watershed genetic variation of coastal cutthroat trout. *Ecological Applications*, **15**, 628-637.
- Yu W, Hardin PE (2006) Circadian oscillators of *Drosophila* and mammals. *Journal of Cell Science*, **119**, 4793-4795.

“A trout,
it swims...

it swims.”

A.H.J.D.