2009

Gene flow & dispersal as factors mediating population divergence, admixture and speciation in fishes

Ryan P. Walter
University of Windsor

GENE FLOW & DISPERSAL AS FACTORS MEDIATING POPULATION DIVERGENCE, ADMIXTURE AND SPECIATION IN FISHES

by

Ryan P Walter

B.Sc., SUNY Brockport, NY, 2000
M.Sc., SUNY Brockport, NY, 2002

A Dissertation
Submitted to the Faculty of Graduate Studies through the
Great Lakes Institute for Environmental Research in partial fulfillment of the
Requirements for the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada
2009

Your file Votre référence
ISBN: 978-0-494-82070-4
Our file Notre référence
ISBN: 978-0-494-82070-4

Declaration of Co-Authorship / Previous Publication

I. Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is result of joint research, as follows: Chapter 2 contains material from a manuscript entitled "Gene flow increases temporal stability of Chinook salmon (*Oncorhynchus tshawytscha*) populations in the Upper Fraser River, British Columbia, Canada" that has been accepted in the Canadian Journal of Fisheries and Aquatic Sciences. This manuscript is co-authored by RP Walter, T Aykanat, DW Kelly, JM Shrimpton and DD Heath. Chapter 3 contains material from the paper entitled "Dispersal and population genetic structure of Telmatherina antoniae, an endemic freshwater Sailfin silverside from Sulawesi, Indonesia" published in The Journal of Evolutionary Biology. This manuscript is co-authored by RP Walter, GD Haffner and DD Heath. Chapter 4 contains material from a manuscript entitled "No barriers to gene flow among sympatric colour-morphs of "small" Telmatherina antoniae from Indonesia's Lake Matano" that has been submitted to The Journal of Fish Biology. This manuscript is coauthored by RP Walter, GD Haffner and DD Heath. The Appendix contains the manuscript "Characterization of four tetranucleotide and six dinucleotide microsatellite markers for use in the tropical freshwater fish *Telmatherina antoniae* and related species" published in Molecular Ecology Notes. This manuscript is co-authored by RP Walter, JR Ovenden and DD Heath. In all cases, the key ideas, primary contributions, experimental designs, data analysis and interpretation, were performed by the author, and the contribution of co-authors was primarily through the provision of guidance with field and lab work.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis. I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publication

This thesis includes four original papers that have been previously published/submitted for publication in peer reviewed journals, as follows:

Thesis Chapter	Publication title/full citation	Publication status*
Chapter 2	Walter RP, Aykanat T, Kelly D, Shrimpton JM, Heath DD. Gene flow increases temporal stability of Chinook salmon (*Oncorhynchus tshawytscha*) populations in the Upper Fraser River, British Columbia, Canada. *Canadian Journal for Fisheries and Aquatic Science*	In press
Chapter 3	Walter RP, GD Haffner GD, Heath DD (2008) Dispersal and population genetic structure of	Published

	Telmatherina antoniae, an endemic freshwater Sailfin silverside from Sulawesi, Indonesia. *Journal of Evolutionary Biology*. DOI: 10.1111/j.1420-9101.2008.01645.x	
Chapter 4	Walter RP, Haffner GD, Heath DD. No barriers to gene flow among sympatric colour morphs of *Telmatherina antoniae* from Indonesia's Lake Matano. In review at *Journal of Fish Biology*.	Submitted
Appendix	Walter RP, Ovenden JR, Heath DD (2007) Characterization of four tetranucleotide and six dinucleotide microsatellite markers for use in the tropical freshwater fish *Telmatherina antoniae* and related species. *Molecular Ecology Notes* 7: 651-653. DOI: 10.1111/j.1471-8286.2006.01664.x	Published

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Understanding how species persist and how they arise are among the most fundamental objectives of ecology and evolutionary biology. Genetic markers coupled with population genetic theory are frequently employed in the study of animal dispersal and gene flow, two parameters central to understanding the genetic structure underlying natural populations.

This thesis presents novel examples of the temporally-stabilizing effect of low to moderate levels of gene flow (Chapter 2), population admixture or evidence of sympatric divergence (Chapter 3), a lack of genetic structure among phenotypes with no reproductive barriers (Chapter 4), river effects on genetic structure of endemic populations (Chapter 5), and parallelism in ecological diversification (Chapter 6).

Throughout this thesis, dispersal also plays a fundamental role in the observed genetic structure, taking on a number of forms given the life history of the species under study. For homing salmonids, straying from natal spawning sites is dispersal. In Chapter 2, I argue that low levels of gene flow can elevate effective population sizes and preserve genetic variability highlighting the importance of considering gene flow acting to temporally stabilize populations. The genetic data presented here overwhelmingly suggests that dispersing individuals are likely reproductively successful members of the populations into which they stray (Chapter 2). For the Malili Sailfin silversides, larval dispersal most likely accounts for the sympatric population structure of an endemic fish in Lake Matano (Chapter 3, 4). Population genetic structure in the presence of substantial dispersal provides a unique perspective on the evolution of reproductive isolation. Genetic analyses using novel microsatellite markers revealed significant population-level

structure consisting of 4-6 sympatric clusters resolving the apparent paradox of population genetic structure coupled with frequent dispersal.

Despite an ontogenetic predisposition to dispersal, geographic and other environmental factors also encourage or retard dispersal. Rivers act to facilitate the movement of individuals between lacustrine populations but biasing one lake as a new exporter of migrants (Chapter 5), or they can completely block passage leading to isolation and the independent evolution of similar feeding strategies (Chapter 6). In all these cases, dispersal (or a lack of) influences the underlying population genetic structure.

For my Grandparents

Bucky and Suzanne

ACKNOWLEDGEMENTS

I thank my primary supervisor Daniel Heath. His breadth of understanding in biology is the unattainable goal I have set for myself throughout my career. I am also grateful for the inspiration and guidance of my co-supervisor Doug Haffner. Doug's ability to see outside the narrow scope of most student projects and to place an idea among the broader picture of progressive science is seriously envied. I also acknowledge the guidance of Melania Cristescu and Daniel Mennill who served on my dissertation committee.

Tutku Aykanat is the first friend I made in Windsor. The many days of discussing evolutionary ideas formed the basis of a new way of thinking about science for me. Tutku is responsible for teaching me the standard methods of molecular techniques.

Jody Shields is among one of the finest young scientists I have ever met. Her enthusiasm towards not only her own work but mine as well, served as a great inspiration. Outside of the laboratory she is a source of inspiration and the love of my life.

Courtney Beneteau, Derek Hogan, David Kelly, Romina Bortoluzzi and Roger Thiessen reviewed many of my drafts and final manuscripts before I sent them to my advisors. Sinyo Rio greatly assisted me in the field in Sulawesi, Indonesia. Pak Dula cheerfully transported me to all my sites on Lake Matano. Andy Bramburger, Sean Crowe, Andrew O'Neill, Peter Hehanussa, Kaimuddhin, Deki, Lili, and Sarah also assisted with logistics in Indonesia.

I also thank my family for their love and support but also for understanding my absence over the years. And lastly, I thank my small canine friends: Amber, Mollie, Mac and Cory who have come into and out of my life along this adventure.

TABLE OF CONTENTS

LIST OF TABLES

Appendix 1.0

LIST OF FIGURES

1.0 INTRODUCTION

Individuals of the same species typically occur in groups within a specific geographic area. In many cases these groups of individuals, through varying degrees of interbreeding, constitute local populations that collectively represent the geographic range of a species. These local populations possess a unique suite of genes, or gene pool, and may share alleles between these pools due to spatial and temporal overlap in habitat, migratory behaviour or movement between habitats, and other behaviours.

With few exceptions, virtually every group of organisms studied to date exhibits some genetic differentiation among populations (Freeland 2005). While genetic differentiation may be measured using a variety of indices, it is essentially defined by differences in the frequency of alleles at a number of prescribed loci within a defined number of individuals (i.e., a population). The basis for allelic variation at a locus is mutation, a change in DNA base pair sequences caused by a number of factors including errors in DNA replication and exposure to radiation or mutagenic chemicals (Futuyuma 1998; Freeland 2005). Over time, the number of allelic variants at a locus within a population increases or decreases via three evolutionary forces: selection, random genetic drift and gene flow. It is the action, interaction, and counteraction of these three evolutionary forces that provides the basis for the theoretical mechanisms of population divergence (Fig 1.1).

1.1 THEORETICAL MECHANISMS OF POPULATION DIVERGENCE

There is a tendency for populations to exhibit some level of genetic and phenotypic dissimilarity with respect to one another—this distinctiveness is the result of selection, drift and gene flow

acting upon genetic variation of the population. Selection acts by directly increasing or decreasing the number of individuals exhibiting certain phenotypic traits. Positive selection works to favour individuals with traits that confer an advantage in survival and reproduction; whereas negative (purifying selection) acts against individuals with traits that do not. Therefore, selection acts not only on a particular individual and its phenotype, but also on the underlying gene complexes and their respective alleles. Selection primarily, but not exclusively, operates within the coding regions of the genome, while variation in both coding and non-coding regions pass through a filter of genetic drift.

Random genetic drift is the change of a population's allele frequencies by chance alone and exists because survival and reproductive success is variable within a population. For simplicity, if we consider non-coding neutral loci, random genetic drift is assumed to be the principal source of genetic divergence among populations. This assumption is made if the effects of other processes—mutation, selection, and gene flow—are negligible over the temporal scale of study (Wang & Whitlock 2003). Random changes in allele frequencies occurring in separate populations are independent of one another, such that allele frequency distributions of two populations do not change in unison; hence they subsequently differ in observed allele frequencies. Drift is always occurring, but as it is random, its effects are often more pronounced in populations with fewer numbers of breeding individuals. The cumulative effects of random drift and selection within a population are sometimes mediated by an exchange of genetic material with another population.

Gene flow carries alleles from one population to another, potentially changing the frequency of alleles already present in the recipient population or introducing new ones. As such, gene flow operates both as an agent for and against divergence. For example, intermittent, small amounts of gene flow from irregular sources can mimic genetic drift (Wang & Whitlock 2003). However, gene flow typically unites two or more gene pools spatially and temporally through a more or less regular exchange of alleles. This exchange is governed not only by an organism's fecundity, but largely through its mobility.

1.2 DISPERSAL

Dispersal is a widely studied phenomenon in ecology as the mobility of an organism or its gametes sets the limits of both a species' range and the degree of gene flow among local populations. However, the relationship between dispersal and gene flow is subtly more complex. Nearly all cases of genetic exchange between populations are the result of successful dispersal. This common strategy is classically interpreted as the result of selection for any phenotype associated with dispersal and the spread of beneficial alleles (Lenormand 2002). Avoidance of general competition due to overcrowding or risk spreading (Oliveiri et al. 1995, Venable & Brown 1988) and competition from kin (Hamilton & May 1977) are among the explanations favouring the evolution of dispersal. However, a shortfall of this interpretation is that dispersal does not equate with gene flow. As gene flow must be preceded by dispersal, if dispersal is unsuccessful or blocked completely, the results for populations include isolation, genetic structure, local adaptation, and speciation.

1.3 Divergence, Local Adaptation & Speciation

Genetic isolation of a population through time can result in marked differences in allele frequencies. Those differences, or genetic divergence, constitute genetic structure within a species, and it is the basis of these differences that underlie the processes of local adaptation and speciation

Local adaptation is defined as the evolution of traits that increase survival and reproduction within population experiencing environmental conditions specific to a given geographic region (Taylor 1991). Despite gene flow among populations, selection for traits maximizing fitness in a local environment leads to local adaptation. Speciation is a process of forming new species; a process of reproductive isolation associated with phenotypic and/or genetic differences conspicuous enough to warrant separate classification. Thus, both local adaptation and speciation are theoretically rooted in the constraint or cessation of gene flow—gene flow which can only occur if reproductively viable members of the same species mate, which does not always occur.

The speciation process, however, is concerned with the formation of reproductive isolation. A central problem in the study of speciation concerns the origin of isolating barriers, particularly those acting to prevent gene flow among sympatric species (Coyne & Orr 2004). The possibility that several isolating barriers may act together to prevent gene flow illustrates the issue of resolving present importance of barriers and historical importance (Coyne & Orr 2004), however barriers to dispersal can generally be classified as prezygotic.

An individual arriving in a new population via dispersal provides the potential for gene flow between the population of origin and the newly entered population (Bohonak 1999). For situations where dispersal occurs, but gene flow is quite limited or does not occur at all between *dispersers* and *residents*; gene flow fails to homogenize alleles from the dispersed individuals into the existing population (Whitlock & McCauley 1999), rather populations are admixed.

1.4 Dispersal & Population admixture

Movement of a number of individuals from one population into a new population may occur. What happens if these individuals do not interbreed with local members of the recipient population, but rather continue breeding with only each other in their new location? The result is a population admixture, a group of individuals (of the same species) possessing unique genetic clusters. Population admixtures are readily apparent in humans where they are referred to as subpopulations or even local communities. Just as social and cultural norms constrain reproduction in humans, natural and sexual selection constrain reproduction in other populations to produce subpopulation structure independent of spatial distribution.

1.5 Fishes as ideal model organisms

Fish as a group are particularly dispersive organisms. Among vertebrates, fishes exhibit a variety of life history traits making them suitable for the study of dispersal. The possession of fins to propel their bodies through the water medium facilitates movement. The production of oil-filled eggs and mobile larval stages serves to spatially isolate juveniles from adult members of the same population. Furthermore, the general high fecundity (i.e., high number of offspring)

observed across fishes as a whole, increases the chance that at least one viable offspring can successfully move into a new environment.

1.6 Study Species & systems

The studies presented here concern two families of fishes: salmonids (Salmonidae) and Sailfin silversides (Telmatherinidae). Species from both families are well suited for comparative studies on the effects of gene flow on population genetic structure by providing examples of species that occupy differing spectrums of the dispersal spectrum.

Pacific salmonids

Pacific salmonids are ideal for the studies involving questions of dispersal or gene flow as many species exhibit philopatry, a phenomenon where individuals reproduce at the same site as their parents. This occurrence of natal homing has resulted in the partial or complete reproductive isolation of some populations (Quinn 1993). Among the well-studied Pacific salmonids is Chinook salmon (*Oncorhynchus tshawytscha*). The anadromous and semelparous nature of this species, along with variation in the age at which they return to reproduce, makes Chinook salmon a suitable study species for testing questions concerning "straying" and its effects on spatial and temporal population structure.

Populations of Chinook salmon in the Upper Fraser River have been impacted to varying degrees by habitat loss, overfishing, and diversion of tributary headwaters (Shrimpton & Heath 2003). However, these populations are also well-studied, with extensive temporal census and genetic data (Shrimpton & Heath 2003). The use of this dataset (Shrimpton & Heath 2003)

incorporating populations that have experienced differing levels (and scales) of environmental perturbations, presents a unique opportunity to test for the presence of temporal instability in population genetic structure. Furthermore, as there is typically movement among populations through "straying" (Stearns & Hendry 2004), this dataset offers an opportunity to test for effects of "straying" on population genetic structure through time.

Sailfin silversides

Sailfin silversides are a family of small fishes endemic to the freshwaters of Sulawesi, Indonesia, and the mangrove swamps of New Guinea (Saeed & Ivanstoff 1991). Recent studies have placed the Telmatherinidae of the Malili Lakes, a system of five lakes in south central Sulawesi, as possibly the best example of a sympatric species radiation in tropical ancient lakes outside of the cichlids (Herder et al. 2006; Roy et al. 2007). These lakes are characterized by clear waters and considered oligotrophic as evidenced by limited biological productivity (Haffner et al. 2006). The headwaters of the system: Lake Matano, has received considerable attention likely attributed to its extraordinary depth (590 m) and age (2-4 Myr, Brooks 1959; Hamilton 1979). A number of other species flocks are found within the system and are the subject of much study: gastropods (von Rintelen et al. 2004), shrimps (Roy et al. 2006) and crabs (Schubart et al. 2008).

Patterns of discordance concerning the endemism of extant taxa inhabiting the Malili Lakes system poses a number of questions colonization modes and routes in relation to the age and formation of the lakes (Roy et al 2004; Herder et al 2006; Schubart et al 2008; von Rintelen et al 2004). For example, Chia & Ng (2006) and Schubart & Koller (2006) have argued that decapods

have at least twice, independently colonized the system. Work in the gastropod *Tylomelania* suggests that this genus has colonized the lake system several times independently (von Rintelen et al. 2004). These patterns in invertebrates are contrasted by the arguments that all extant endemic species of telmatherinids of Lake Matano have descended one single colonization event (Roy et al. 2007; Schubart et al. 2008) and evidence for multiple introgressive hybridization events (Herder et al. 2006). Nonetheless, this discrepancy across taxa is likely the result of ontogenetic differences in colonizing ability, but is nonetheless an important point which remains unclear despite a number of studies classifying the divergent patterns in this family as an example of adaptive radiation (Roy et al. 2004; Herder et al. 2006).

Within the Telmatherinidae of the Malili lakes are three main genera: *Telmatherina*, *Paratherina*, and *Tominanga*. All species of *Telmatherina*, save *T. bonti*, are endemic to Lake Matano. *Paratherina* and *Tominanga* are distributed throughout Lakes Mahalona and Towuti, but not found in Matano. The isolation of species and genera in this system raises questions surrounding the nature of dispersal (or a lack thereof) between lakes. Recent limnological work suggests varying degrees of metal toxicity and nutrient availability (Crowe et al. 2008, Sabo et al. 2008) which may confer differential physical and biological pressures within this system.

1.7 THESIS OBJECTIVES

How can individuals within divergent populations remain members of the same species? The answer is gene flow. Therefore, resolving the level of isolation and gene flow among populations remains central to the concepts of local adaptation and speciation. Using freshwater and anadromous fishes as model organisms, I demonstrate how dispersal mediates genetic divergence

to result in differentially structured populations. The majority of the genetic data included here consist of multilocus microsatellite genotyping and the application of individual-based assignment tests. These techniques are standard tools in molecular ecology and widely believed to hold the potential to estimate rates of dispersal and gene flow (Waser & Strobeck 1998; Berry et al. 2004).

Chapter 2 examines the temporal genetic stability of five populations of Chinook salmon (*Oncorhynchus tshawytscha*) from the upper Fraser River, British Columbia, Canada. Using population genetic analysis of microsatellite loci, I characterize the spatial and temporal genetic structure of these populations. Using genotype assignment techniques to identify dispersed individuals, I demonstrate that these individual fish are breeding members of the populations (i.e., true migrants) in which they were sampled and that they confer increased temporal genetic stability to their recipient population. This chapter provides a unique empirical example of gene flow working to stabilize populations through time. As low straying levels in salmonids may have evolved to favour local adaptation, I demonstrate in stark contrast, that even such low levels of gene flow can elevate effective population sizes yet still preserve spatial genetic variability. This chapter highlights the importance of considering how gene flow acts to temporally stabilize populations, particularly small populations where effects of interruptions in more or less 'regular' gene flow will be most pronounced.

Chapter 3 provides a unique perspective on the evolution of reproductive isolation in the presence of substantial dispersal. In this chapter I examine the population genetic structure of an endemic fish, *Telmatherina antoniae*, throughout its known range; a single lake in Indonesia. I

provide evidence of the existence of at least four sympatric populations and an ability of this fish to disperse throughout its range within one generation. This chapter highlights the importance of considering cryptic genetic structure within a species, and resolves an apparent paradox of population genetic structure coupled with frequent dispersal.

Chapter 4 further examines the persistence of colour polymorphism within Lake Matano's *Telmatherina antoniae*. I assess genetic divergence and inter-male aggression among sympatric colour morphs on both a 'site-by-site' and 'lake-wide' level. This chapter provides both genetic and behavioural perspectives on the maintenance of colour polymorphisms in *T. antoniae* despite no barriers to gene flow.

Chapter 5 demonstrates the effects of directional gene flow on population connectivity. Using two species of *Paratherina* from Sulawesi's southern Malili Lakes, I show that asymmetrical gene flow can intensify genetic drift in source populations. These results provide important considerations for the conservation and management of source populations in terms of genetic diversity and a viable fishery.

While dispersal appears to be evident among the Telmatherinid fishes of the Malili Lakes, barriers to both dispersal and gene flow do exist within the system and are likely responsible for the observed endemism of particular genera. Chapter 6 demonstrates how barriers to dispersal and gene flow have resulted in the independent evolution of similar ecological strategies in separate lakes. Uniting new data on the adaptive divergence of *Paratherina* species with published data of the closely-related *Telmatherina*, I provide an example of parallel

ecological divergence which strengthens the argument that the diversity observed in the Telmatherinid radiation is primarily the result of ecological speciation.

In summary, I provide three empirical examples of the differential contributions of dispersal on gene flow and population genetic structure in fishes. Additionally, I also present an example of how a lack of gene flow and isolation of closely-related taxa experiencing similar ecological pressures can produce very similar evolutionary patterns.

1.8 REFERENCES

Berry O, Tocher MD, Sarre SD 2004. Can assignment tests measure dispersal? *Mol Ecol* 13, 551-561.

Bertram J 2000. The molecular biology of cancer. *Mol Aspects Med.* 21, 167–223.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. *Q. Rev Biol* **74**: 2145.

Brooks JL 1950. Speciation in ancient lakes concluded. *Q Rev Biol* 25, 131-176.

Chia OKS, Ng PKL 2006. The freshwater crabs of Sulawesi, with descriptions of two new genera and four new species (Crustacea: Decapoda: Brachyura: Parathelphusidae). *Raffles Bull Zool* 54, 381-428.

Coyne JA, Orr HA 2004. Speciation. Sinauer, Sunderland, MA. 545 pp.

Crowe SA, O'Neill AH, Katsev S , Hehanussa P, Haffner GD, Sundby B, Mucci A, Fowle D 2008. The biogeochemistry of tropical lakes: A case study from Lake Matano, Indonesia. *Limnol Oceanogr* 53, 319–331.

Freeland JR 2005. Molecular Ecology. Wiley & Sons, NJ. 388 pp.

Futuyuma DJ 1998. Evolutionary biology. 2d ed. Sinauer, Sunderland, MA. 763 pp.

Hamilton WD, May RM 1977. Dispersal in stable habitats. *Nature* 269, 578–581.

Hamilton W 1979. *Tectonics of the Indonesian region.* Geological survey professional paper 1078. Washington, United States Government Printing Office.

Herder F, Nolte AW, Pfaender J, Schwarzer J, Hadiaty RK, Schliewen UK 2006. Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc Roy Soc B* 273, 2209-2217

Lenormand T 2002. Gene flow and the limits to natural selection. *TREE* 17. 183-189.

Oliveiri I, Michalakis Y, Gouyon PH 1995. Metapopulation genetics and the evolution of dispersal. *Am Nat* 146, 202-228.

Quinn TP 1993. A review of homing and straying of wild and hatchery-produced salmon. *Fish Res* 18, 29-44.

Rintelen T von, Wilson AB, Meyer A, Glaubrecht M 2004. Escalation and trophic specialization drive adaptive radiation of freshwater gastropods in ancient lakes on Sulawesi, Indonesia. *Proc Roy Soc B* 271, 2541-2549.

Roy D, Docker MF, Hehanussa P, Heath DD, Haffner GD 2004. Genetic and morphological data supporting the hypothesis of adaptive radiation in the endemic fish of Lake Matano. *J Evol Biol* 17, 1268-1276.

Roy D, Kelly DW, Fransen CHJM, Heath DD, Haffner GD 2006. Evidence of small-scale vicariance in *Caridina lanceolata* (Decapoda: Atyidae) from the Malili Lakes, Sulawesi. *Evol Ecol Res* 8, 1087-1099.

Roy D, Paterson G, Hamilton PB, Heath DD, Haffner GD.2007. Resource-based adaptive divergence of the freshwater fish Telmatherina from Lake Matano, Indonesia. *Mol Ecol* 16, 35-48.

Sabo E, Roy D, Hamilton PB, Heahanussa PE, McNeelyR, Haffner GD 2008. The plankton community of Lake Matano: factors regulating plankton composition and relative abundance in an ancient, tropical lake of Indonesia. *Hydrobiologia* 615:225–235

Saeed B, Ivantsoff W 1991. Kalyptatherina, the first telmatherinid genus known outside of Sulawesi. *Ichthyol Explor Freshw* 2, 227–238.

Schubart CD, Koller P 2006. Multiple colonization of ancient lakes in Sulawesi by arathelphusid crabs (Crustacea: Brachyura). *In:* The Ecology and Limnology of the Malili Lakes. Hehanussa PE, Haryani GS and Ridwansyah I, eds. Proceedings International Symposium, Bogor, Indonesia 20-22 March 2006, pages 47-53. LIPI Press, Jakarta, Indonesia. ISBN 979-26-2407-4.

Schubart CD, Santl T, Koller P 2008. Mitochondrial patterns of intra- and interspecific differentiation among endemic freshwater crabs of ancient lakes in Sulawesi. *Contrib Zool* 77, 83-90.

Shrimpton JM, Heath DD 2003. Census vs. effective population size in Chinook salmon: large- and small-scale environmental perturbation effects. *Mol Ecol* 12, 2571-2583.

Stearns SC, Hendry AP 2004.The salmonid contribution to key issues in evolution. *In* Evolution Illuminated. *Edited by* A.P Hendry and S.C. Stearns. Oxford University Press, New York, USA. pp 53-90.

Taylor EB 1991. A review of *local adaptation* in Salmonidae, with particular reference to Atlantic and Pacific salmon. *Aquaculture* 98, 185-207.

Venable DL, Brown JS 1988. The selective interaction of dispersal, dormancy, and seed size as adaptations for reducing risk in variable environments. *Am Nat* 131, 360–384.

Wang J, Whitlock MC 2003. Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* 163, 429-446.

Waser PM, Strobeck C 1998. Genetic signatures of interpopulation dispersal. *TREE* 13, 43-44.

Whitlock MC, McCauley DE 1999. Indirect measures of gene flow and migration: F_{ST} doesn't equal 1/(4Nm+1). *Heredity* 82, 117-125

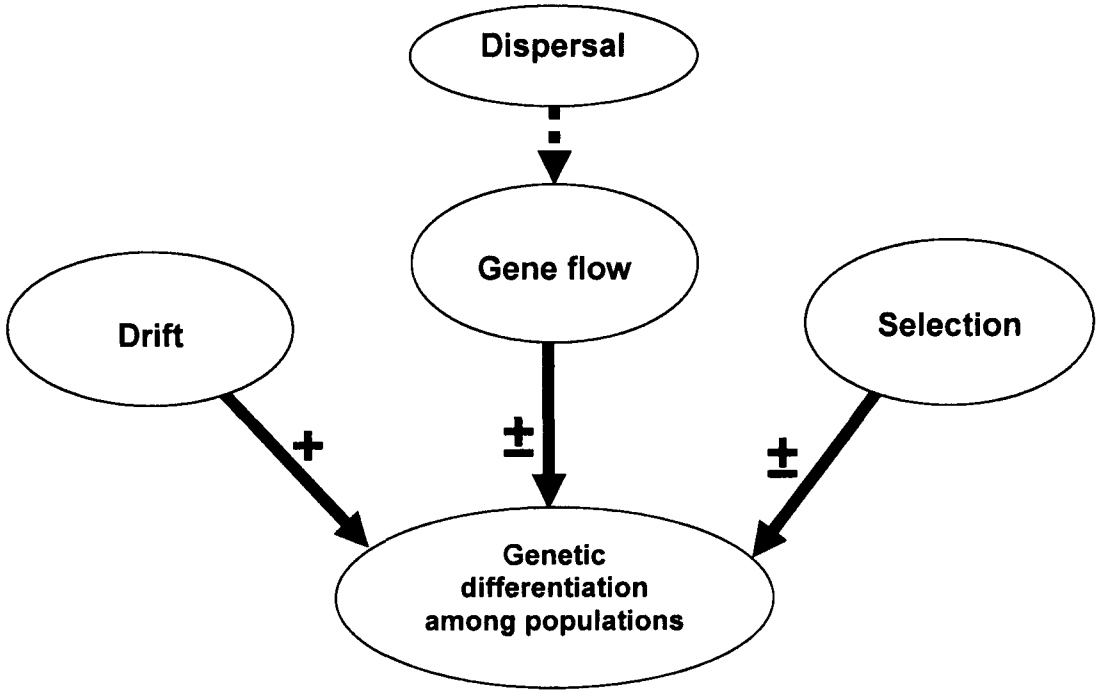


Figure 1.1. A schematic diagram showing dispersal and its influence on gene flow in conjunction with other evolutionary forces that control population genetic divergence. Drift works to increase differentiation by randomly driving alleles in two isolated populations either to fixation or extinction. Selection can promote genetic differentiation by favouring different alleles if those alleles promote increased survival and reproduction specific to pressures exerted in each population. Selection can also oppose genetic differentiation through balancing or stabilizing selection. Gene flow opposes population differences by sharing alleles through interbreeding, directly opposing the differentiation imposed by drift and selection. However, gene flow can itself also promote genetic differentiation if gene flow is intermittent in source and rate, which may introduce new alleles, promoting the creation of novel genotypes.

2.0 — GENE FLOW INCREASES TEMPORAL STABILITY OF CHINOOK SALMON (*ONCORHYNCHUS TSHAWYTSCHA*) POPULATIONS IN THE UPPER FRASER RIVER, BRITISH COLUMBIA, CANADA*

2.1 INTRODUCTION

Analyses of population genetic structure using microsatellite markers have widespread application in evolutionary and conservation biology. Those applications include the identification of populations of specific conservation concern, such as those severely bottlenecked or otherwise displaying reduced genetic diversity. However, the supposition of little or no recent genetic change (i.e., genetic stability) can be misleading as allele frequencies may fluctuate temporally as a consequence of variation in factors such as effective population size (N_e), species range and dispersal, and life history. For example, random drift based genetic change may confound conservation-based analyses testing for changes in genetic composition resulting from anthropogenic or environmental effects on population processes (Heath et al. 2002). Temporal genetic samples are available for a number of aquatic species due to their high abundance in nature and the availability of archived material. This is especially true of commercially important fish species, including cod (Ruzzante et al. 1997; Poulsen et al. 2006), sea basses (Rhodes et al. 2004), and various salmonids (Østergaard et al. 2003; Palm et al. 2003; Jensen et al. 2005). Among salmonids, there is no consistent pattern of genetic stability, as some studies have reported stable genetic structure and composition (Tessier and Bernatchez 1999; Heath et al. 2002; Hansen et al. 2002) whereas a number of papers report genetic instability (Østergaard et al. 2003; Palm et al. 2003; Jensen et al 2005). As more studies are published with

*Walter RP, Aykanat T, Kelly DW, Shrimpton JM, Heath DD Gene flow increases temporal stability of Chinook salmon (*Oncorhynchus tshawytscha*) populations in the Upper Fraser River, British Columbia, Canada Accepted by *Canadian Journal of Fisheries and Aquatic Sciences* Oct 10, 2008

estimates of temporal genetic change, the potential for identifying the factors affecting temporal stability will likely increase.

Previous work in salmonids has demonstrated that, theoretically, significant variation in spatial and temporal genetic structure may result from interactions between drift and local adaptation (Elo 1993; Adkinson 1995). While variation at gene loci is driven by selection, mutation, drift, and gene flow, small populations will encounter elevated drift effects and may appear to be less stable over time due to stochastic effects alone. This elevated variance may be tempered by intermittent gene flow serving to inflate local N_e (Slatkin 1985). For example, gene flow between life history stages within a single river has been shown to have a stabilizing effect on N_e (Araki et al. 2007). Alternatively, episodic gene flow among genetically divergent populations may artificially inflate allele frequency variance, increasing apparent genetic instability and leading to low estimates of N_e. With this in mind, relatively small populations that experience gene flow are ideal models with which to assess factors that affect temporal genetic stability, since genetic drift is likely to be high and the effect of gene flow is likely to be profound.

Pacific salmonids are particularly useful species in which to examine the relationship between gene flow and genetic stability since high levels of natal philopatry result in low rates of straying (Altukhov et al. 2000). The resulting limited gene flow among even geographically close populations often leads to differentiation via genetic drift and local adaptation (Taylor 1991; Quinn 1993; Heath et al. 2006). Indeed, a number of studies using microsatellite loci show

particularly high levels of genetic structure in salmonid populations (e.g. Taylor 1995; Nielsen et al. 1997).

I use an historical microsatellite marker data set (1978 – 1998) from Shrimpton and Heath (2003) to test novel hypotheses concerning the temporal stability of allele frequencies in five populations of Chinook salmon (*Oncorhynchus tshawytscha*) in the Upper Fraser River, BC Canada. Chinook salmon were sampled at approximately 10-year intervals from the late 1970s to the late 1990s, and those samples proved useful for the determination of N_e (Shrimpton and Heath 2003). Here, I first quantify temporal changes in genetic structure and allele frequency distribution to test the hypothesis that populations are temporally stable. I then examine the relationship between migration and temporal stability to test the hypothesis that gene flow through time reduces temporal allelic variance and raises N_e, thereby lending to increased population viability in the long term.

2.2 Materials and Methods

Study areas

The five study populations for which microsatellite genotype data from Shrimpton and Heath (2003) were available are shown in Figure 2.1. The Nechako River (N), Stuart River (S), Willow River (W), Bowron River (B), and Dome Creek (D) populations were chosen based on the availability of archived scale samples, close geographical proximity, and well-documented historical information. Within the last 30 years, census population sizes have either increased from moderate (Bowron) or very low numbers (Willow), remained small (Dome), remained relatively stable over time (Nechako) or shown considerable annual variation (Stuart; Shrimpton

and Heath 2003). For each river, scale samples of adult spawners were collected at two or three dates over a 20-year period (1978–98). Sample size for each population and sampling time ranged from 29 to 60 individuals, except Bowron 1998, which consisted of 18 individuals. Allele frequency data was calculated for dinucleotide repeat markers (Ots1, Ots3, Ots4, Omy207. Omy325, Oneu3, Ssa85, Ssa197, Sfo8) and tetranucleotide repeat markers, (Ots104, Ots107). Individuals with genotypes missing at two or more loci, and those with rare alleles (< 0.01, study-wide across all time periods) were removed from the data set prior to analyses.

Statistical analysis

The 11 microsatellite loci used here had high levels of variation in the Chinook salmon populations examined, while nine tests out of 154 showed significant deviation from Hardy-Weinberg equilibrium, but without a consistent pattern. To test for differences in allele frequency distribution among sampling years and populations I performed pairwise exact tests using **TFPGA 1.3** (Miller, 1997; 1000 dememorization steps; 20000 permutations; Raymond and Rousset 1995). To quantify genetic divergence among populations and sampling years, I calculated pairwise F_{ST} (Weir and Cockerham 1984) and estimated significance with 10000 permutations using **MSA** software (Dieringer and Schlötterer 2003). I applied Bonferroni correction to maintain an experiment-wide α-level of 0.05 for all pairwise comparisons. Principal Coordinates Analysis (PCoA, covariance-standardized) was performed on the pairwise matrix of F_{ST} values in **GENALEX 6.1** (Peakall and Smouse 2007) to show relationships among samples.

To partition allele frequency variance among populations, among sampling years within populations, and among individuals, I performed a hierarchical analysis of molecular variance

(**AMOVA**, Excoffier et al. 1992) using **ARLEQUIN** software (Schneider et al. 2000). Significance values of the variance components were obtained using 10000 permutations.

To quantify the degree of temporal allelic variation within each population, I estimated the standardized allele variance ($\hat{F}$, Waples 1989) and effective population size (N_e, Waples 1989) using **N_e ESTIMATOR** (Peel et al. 2004). Estimates of N_e were corrected for overlapping generations by multiplying N_e estimates by mean generation times (Shrimpton and Heath 2003). Both estimates were calculated for each pairwise temporal comparison within populations.

To determine the influence of migration on temporal stability I identified first-generation migrants based on population exclusion methods in **GENECLASS 2.0** (Piry *et al.* 2004) using the Bayesian method (Rannala and Mountain 1997) and likelihood ratio *L-home / L-max* (Paetkau et al. 2004). Monte-Carlo re-sampling was performed with 10000 simulated individuals at an assignment threshold p-value of 0.05. I compared D_{LR} values and F_{ST} estimates among populations to determine whether detection of migrants is possible (Paetkau *et al.* 2004). I then recalculated allelic variance ($\hat{F}$) and moments-based N_e in all populations with the identified migrants removed. To determine whether a reduction in sample size resulted in a change in $\hat{F}$ following removal of migrants, I randomly removed individuals (equal numbers as the identified migrants) from each temporal sample in ten replicates, then recalculated $\hat{F}$ (Fig A1).

To further illustrate the influence of migration on temporal stability I also calculated N_e via two methods using the maximum-likelihood approach of Wang and Whitlock (2003) in the software **MNE 2.0**. N_e is inversely proportional to drift, and hence temporal allelic variance. To

determine the effect of migrant fish on N_e, I first estimated N_e assuming 'closed' populations on the original dataset, and then re-ran the analysis following removal of the migrants identified by **GENECLASS**. Next I estimated N_e allowing migration (e.g., 'open' model) between focal populations (T_1, T_2, T_3 for Nechako, Stuart, Bowron, Dome; and T_1, T_2, for Willow) and a source population. Source populations consisted of pooled genotypes from all samples and time periods except the focal population, a technique which is suggested to be robust (Wang and Whitlock 2003). Generation times of 1 were used for temporal samples separated by 5 years; 2 for 8, 10, and 12 years; 4 for 20 years. N_e estimates were corrected for overlapping generations using the mean generation times as in Shrimpton and Heath (2003). I then compared estimates of all three N_e estimates and their respective N_e / N_c ratios to see whether trends were consistent across methods.

To determine directionality of gene flow among populations, I obtained estimates of migration rate (m) using **BAYESASS 1.2** software (Wilson and Rannala 2003). Runs were performed allowing migration among samples in 3 temporally relevant groupings: 1) N-78, S-80, B-80, W-80; 2) N-88, S-88, B-88, D-86; and 3) N-98, S-98, B-98, D96. Estimates of m were taken from averages of three replicate runs of 5×10^7 iterations, with 3×10^7 burn-in runs, and five subsequent replicate runs of 3×10^7 to identify convergence among runs.

2.3 RESULTS

Exact tests revealed significant differentiation in population genetic structure for all pairwise comparisons ($P < 0.001$) following Bonferonni correction. All pairwise F_{ST} estimates were significant following Bonferonni correction with only four exceptions, only one of which was a

spatial comparison (Table 2.1). The PCoA plot of genetic differentiation among populations illustrates the magnitude of temporal variation for some populations compared to others (Fig. 2.2). For example, the proximity of data points for temporal samples from the Nechako and Stuart compared to those of the Dome or Bowron. The **AMOVA** showed significant genetic differentiation both among sampling years, within populations (1.63%, $P < 0.001$), and among populations (1.21%, $P < 0.005$).

Of the 590 individuals in the dataset, 76, or roughly 13%, were identified as migrants based on **GENECLASS** analysis, with Dome Creek harbouring the highest number of identified migrants and the Stuart showing the fewest (Table 2.2). D_{LR} and F_{ST} measurements from pairwise population comparisons showed sufficient discriminatory power was available to detect migrants (Tables 2.1, 2.3). The subsequent exclusion of migrant individuals from our samples amounted to a removal of an average of 5 individuals per site study-wide, resulting in the retention of approximately 87% of all individuals across all samples (Table 2.3). Following removal of migrants, the average sample size was 37 individuals per site. As predicted, increases in allelic variance were noted for all populations for all temporal samples following removal of migrants (Fig 2.3a). Moments-based estimates of N_e showed similar results in that N_e decreased in all populations save the early comparison from the Bowron (Fig 2.3b). Recalculation of $\hat{F}$ following random removal of individuals from each temporal sample (numbers equal to the number of migrants identified) did not yield changes in $\hat{F}$ comparable to those following migrant removal, thus the effect of removing migrants is not a result of the reduction in sample size (Fig. 2.5).

Calculations of N_e using the Wang and Whitlock (2003) methods showed consistently higher N_e when migrants were included in the samples (Fig 2.4a). As N_e estimates are less sensitive to sample sizes, the use of N_e and its subsequent decrease following migrant removal exemplifies the stabilizing effect of the migrants across the study system, independent of sample size effects. Furthermore, the estimates of N_e using the 'open' model were comparable to those using the 'closed' model with migrants removed (Fig. 2.4 a, b). Comparison of N_e / N_c ratios show that this pattern is evident independent of census population size (Fig. 2.4b).

Estimation of migrations rates from **BAYESASS** showed mean migration rates across all sampling periods of 0.032 for Nechako, 0.099 for Bowron, and 0.064 for Stuart. Migration rates estimates for the Dome 1986-88 suggested a possible local optimum trap in **BAYESASS** at 0.33 migration rate, most likely due to a violation of the assumption of < 1/3 migrants (Wilson and Rannala 2003; Austin et al. 2004; Hansen et al. 2007). Work by both Faubet et al. (2007) and Palstra et al. (2007) demonstrate the limitations of **BAYESASS** for estimating recent migration rates, therefore the following migration estimates should be interpreted with caution. **BAYESASS** indicated asymmetrical gene flow (i.e., non-overlap of confidence intervals) from the Nechako into the Stuart in the temporal sampling period 1978-88 (Table 2.2). However, all other migration rates appeared to be bidirectional due to overlap in confidence intervals, despite some suggestion of asymmetrical migration again from Nechako into the Stuart for the 1998 samples (Table 2.2). For the populations and temporal periods in which **BAYESASS** was able to reliably estimate migration, both Dome (0.0106) and Nechako (1978: 0.0133; 1998: 0.0229) showed the lowest incoming rates were smaller. Across all time periods the Bowron received the most

migrants, and within the 1986-88 sampling period, the Dome appeared to be a source of migrants for both the Nechako and Bowron populations (Table 2.2).

2.4 DISCUSSION

Temporal instability in allele frequency distribution was found for all populations of Chinook salmon sampled from tributaries to the Upper Fraser River, with the overall magnitude of temporal within-population variation exceeding that of among-population variation. The level of stability of some populations (Nechako, Stuart) is consistently higher than others (Willow, Dome, Bowron). The temporal predictability of genetic structure among populations varied from nearly wholly consistent to essentially random, which may explain some of the inconsistencies among previous temporal studies of genetic structure in salmonids. Earlier work has reported both stable (Tessier and Bernatchez 1999; Hansen et al. 2002, Heath et al. 2002) and unstable (Laikre et al. 2002, Østergaard et al. 2003, Jensen et al. 2005) population genetic composition and structure in salmonids. Our results indicate that genetic stability clearly varies among populations of Chinook salmon, even within the restricted geographical range of the Upper Fraser River.

Three possibilities may account for the variability in population genetic stability among the geographically close Chinook salmon populations of the Upper Fraser River. First, within-season variation in the genetic composition of the spawning run coupled with changes in the sampling time could drive temporal instability. Fillatre et al. (2003) showed that comparisons of early and late-run Sockeye salmon (*Oncorhynchus nerka*) accounted for twice the genetic variation of that observed among years within the same river. However, I know of no records of multiple

spawning run structure for Chinook in our study populations, but cannot exclude the possibility of within-run temporal variation, as samples were not collected for specific periods within the annual yearly spawning runs. A second source of variation in temporal instability may stem from differential survival or reproductive success among populations. This point is typically well illustrated by comparing N_e to N_c. While the decoupling of N_e and N_c was shown for these populations by Shrimpton and Heath (2003); and is also observed in other salmonids (Araki et al. 2007; Fraser et al. 2007b), our estimates of N_e (Fig. 2.4) show the persistence of these patterns despite any buffering effects gene flow appears to confer. While this decoupling may suggest that differential survival and reproduction appears independent of gene flow, the consistent declines in N_e for each population following migrant removal/correction further demonstrate that gene flow is the third possibility driving temporal variation in these populations.

Generally, gene flow is expected to be stabilizing if the rate is constant in magnitude, time and source (Wang and Whitlock 2003). Our analyses provide empirical evidence that variation in gene flow is likely driving the differences in temporal genetic stability among our sampled populations. While I conclude that gene flow appears to confer stability for some populations, the relationship between temporal stability, gene flow and neutral genetic differentiation is perhaps more complex as the effects of gene flow on genetic stability clearly depend on not only the source, but also on the genetic stability of the source population itself.

Philopatry in salmon may have evolved to match locally adapted fish with sites exhibiting similarities in habitat or spawning areas (e.g., Stuart and Nechako) suggesting that genetic stability of a recipient population likely depends on the drift-based variance of the source

population. The Nechako and Stuart populations exhibited the greatest temporal stability, but the Stuart shows migration rates from outside are nearly ten-fold higher than in the Nechako. The source of migrants for the Stuart is consistently the Nechako over the study-period, itself a temporally stable system. Such patterns of gene flow are likely the result of straying among the geographically close upper Fraser River Chinook salmon populations, as individuals may find it more difficult to distinguish between proximal natal and non-natal sites (Hendry et al. 2004). This dispersal or straying may not reflect individual "choice" (Hendry et al. 2004) since abiotic factors such as environmental perturbation possibly affect straying rates and patterns (e.g., controlled water discharge in the Nechako River may influence straying into the adjacent, Stuart River). Furthermore, both habitat degradation and population declines have been noted for these populations (Shrimpton and Heath 2003) which likely impedes assessment of natural or original connectivity patterns (Palstra et al. 2007).

Recent work in salmonids has used patterns of asymmetrical gene flow from large into small populations as evidence for density-dependent dispersal (Fraser et al. 2004; Hansen et al. 2007). This pattern is not observed for our populations; however directionality of gene flow has been shown to be primarily determined by the temporal scale over which migration was being estimated (Palstra et al. 2007). As gene flow was estimated indirectly from point-samples between rivers at different times, it is likely all source populations were not sampled and the estimates of specific directional migration rates between populations should be interpreted with caution. The migration rates estimated in **BAYESASS** further illustrate the complexity of the relationship between gene flow and stability as high migration estimates for Bowron 1978 (Table 2.2) do not translate into higher N_e estimates. This trend is also illustrated by the slight increase

in N_e 'open' estimates for the Bowron following removal of migrants using both Moments-based and Wang and Whitlock's methods (Fig. 2.2 b, c). Although there is a possibility that identified migrants may not actually reproduce, the effect of correcting for migration on our estimates of $\hat{F}$ and N_e consistently demonstrates the temporally stabilizing potential of gene flow.

A number of studies show gene flow-induced genetic stability in other species, perhaps hinting at the generality of this effect. Dispersal (and subsequent gene flow) has been credited for the high temporal genetic stability of coyote populations despite aggressive removal efforts (Williams et al. 2003). Studies in insects have also linked temporal stability to consistent gene flow (Loxdale and Brookes 1990, Bourguet et al. 2000). For species exhibiting metapopulation-type structuring, such as cod (*Gadus morhua*), high migration rates buffer against bottlenecks through inflating N_e, as evidenced by drastic reductions in census sizes but no corresponding changes in N_e (Ruzzante et al. 1997. Poulsen et al. 2006).

This study provides a valuable assessment of Wang and Whitlock's (2003) 'open' model for estimating N_e. This model simulates migration between potential source populations and the focal populations at given time periods, thereby N_e estimates using the 'open' model should resemble those of the 'closed' model following migrant removal. For Atlantic salmon, Fraser et al. (2007a) identified N_e 'open' estimates that more closely reflected biological reality than higher N_e 'closed' estimates in populations that were known to have experienced bottlenecks. Fraser et al. (2007a) and Palstra and Ruzzante (2008) further suggest that N_e 'open' estimates may be generally biased downwards and N_e 'closed' estimates biased upwards. Palstra and Ruzzante (2008) observed upward biases of N_e under N_e 'closed' conditions in 94% of their

salmonid estimates signifying that relatively continuous gene flow reduces drift. In this study, the congruence of the N_e 'open' estimates to those of the N_e 'closed' with migrants removed, emphasizes this point that N_e is augmented in Chinook salmon populations, and that these populations are not as isolated as previously believed.

The effect of gene flow on genetic stability has implications for the management and ecology of Chinook salmon. First, the implicit assumption that straying in Pacific salmon is rare (and of little ecological relevance for long-term population viability) should be re-evaluated. As is well known, populations with low N_e must be accorded high conservation priority; however, this study shows that even populations with moderately high N_e may be unexpectedly at risk if migrant source populations are impacted. The results of this study thus emphasize the importance of nearby source populations, particularly those that are genetically stable, as they affect local population stability and help to maintain connectivity and long term population viability. Hence, the definition of management and conservation units should consider temporal stability (and thus N_e) as an important population parameter since it reflects both within-population factors as well as among-population gene flow (Fraser et al. 2007a).

The results of this study are consistent with the hypothesis that populations experiencing regular gene flow may be buffered from temporal genetic variation, resulting in elevated estimates of N_e. From a management perspective, these findings emphasize the value of temporal genetic screening of populations for defining genetic relationships among populations and highlight the importance of gene flow in stabilizing both small and larger populations. As isolation increases genetic instability, isolated populations should be of high management

priority Furthermore, source populations within this system are also of high priority as they contribute to the genetic diversity of receiving populations. Given the range of temporal stability among our study populations, it is clear that no assumptions can be made concerning temporal genetic stability among even geographically close populations of a single species, much less across species.

2.5 REFERENCES

Araki, H., Waples, R.S., Ardren, W.R., Cooper, B., and Blouin, M.S. 2007 Effective population size of steelhead trout: influence of variance in reproductive success, hatchery programs, and genetic compensation between life-history forms. Mol. Ecol. **16**: 953–966.

Adkinson, M. 1995. Population differentiation in Pacific salmon: Local adaptation, genetic drift or the environment? Can. J. Fish. Aquat. Sci. **52**: 2762-2777

Altukhov, Y.P., Salmenkova, E.A. and Omelchenko, V.T 2000. Salmonid fishes: Population biology, genetics and management. Blackwell Science, Oxford.

Austin, J.D., Lougheed, S.C., and Boag, P.T. 2004. Controlling for the effects of history and nonequilibrium conditions in gene flow estimates in northern bullfrog (*Rana catesbeiana*) populations. Genetics, **168**: 1491–1506.

Bourguet, D., Bethenod, M.T., Pasteur, N., and Viard, F 2000. Gene flow in the European cornborer *Ostrinia nubilasis*: implication for the sustainability of transgenic insecticidal maize. Proc. R. Soc. London .B. **167**: 267-122.

Dieringer, D., and Schlötterer, C. 2003. Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes **3**: 167-169.

Elo, K. 1993. Gene flow and conservation of genetic variation in anadromous Atlantic salmon (*Salmo salar*) Hereditas, **119**: 149-159.

Excoffier, L., Smouse, P., and Quattro, J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. Genetics, **131**: 479–491.

Fillatre, E.K., Etherton, P., and Heath D.D. 2003. Bimodal run distribution in a northern population of sockeye salmon (*Oncorhynchus nerka*) life history and genetic analysis on a temporal scale. Mol. Ecol. **12**: 1793-1805.

Fraser, D.J., Lippe, C., and Bernatchez, L. 2004. Consequences of unequal population size, asymmetric gene flow and sex-biased dispersal on population structure in brook charr (*Salvelinus fontinalis*). Mol. Ecol. **13**: 67-80.

Fraser, D.J., Hansen, M.M., Tessier, N., Ostergaard, S., Legault, M., and Bernatchez, L. 2007a. Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. Mol. Ecol. **16**: 3866-3889.

Fraser, D.J., Jones, M.W., McParland, T.L., and Hutchings, J.A. 2007b. Loss of historical immigration and the unsuccessful rehabilitation of extirpated salmon populations. Conserv. Gen. **8**: 527-546.

Hansen, M.M., Ruzzante, D.E., Nielsen, E.E., Bekkevold, D., Mensberg, K-L.D. 2002. Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. Mol. Ecol. **11**: 2523-2535.

Hansen, M.M., Skaala, O., Jensen, L.F., Bekkevold, D., and Mensberg, K.L.D. 2007 Gene flow, effective population size and selection at major histocompatibility complex genes: brown trout in the Hardanger Fjord, Norway. Mol. Ecol. **16**: 1413-1425.

Heath, D.D., Busch, C., Kelly, J., Atagi, D.Y 2002. Temporal change in genetic structure and effective population size in Steelhead trout (*Oncorhynchus mykiss*). Mol. Ecol. **11**: 197-214.

Heath D.D., Shrimpton, J.M., Hepburn, R.I., Jamieson, S.K., Brode, S.K., Docker, M.F.2006. Population structure and divergence using microsatellite and gene locus markers in Chinook salmon (*Oncorhynchus tshawytscha*) populations. Can. J. Fish. Aquat. Sci. **63**: 1370-1383.

Hendry, A.P., Castric, V., Kinnison, M.T., Quinn, T.P 2004. The evolution of philopatry and dispersal: Homing versus straying in salmonids. *In* Evolution Illuminated. *Edited by* A.P. Hendry and S.C. Stearns. Oxford University Press, New York, USA. pp 53-90.

Jensen, L.F., Hansen, M.M., Carlsson, J., Loeschcke, V., and Mensberg, K-L.D. 2005. Spatial and temporal genetic differentiation and effective population size of brown trout (*Salmo trutta*, L.) in small Danish Rivers. Conserv. Gen. **6**: 615-621.

Laikre, L., Jarvi, T., Johansson, L., Palm, S., Rubin, J.F., Glimsater, C.E., Landergren, P., and Ryman, N. 2002. Spatial and temporal population structure of sea trout at the Island of Gotland, Sweden, delineated from mitochondrial DNA. J. Fish. Biol. **60**: 49–71.

Loxdale, H.D., and Brookes, C.P 1990. Temporal genetic stability within and restricted migration (gene flow) between local populations of the Blackberry-grain aphid *Sitobion fragariae* in Southeast England. J. Anim. Ecol. **59**: 497-514.

Miller, M.P 1997. Tools for population genetic analysis (TFPGA) ver. 1.3. Department of Biological Sciences, Northern Arizona University, Flagstaff, Ariz.

Nielsen, J.L., Carpanzano, C., Fountain, M.C., and Gan, C.A. 1997. Mitochondrial DNA and nuclear microsatellite diversity in hatchery and wild *Oncorhynchus mykiss* from freshwater habitats in southern California. Trans. Am. Fish. Soc. **126**: 397-417.

Østergaard, S., Hansen, M.M., Loeschke, V., and Nielsen, E.E. 2003. Long-term temporal changes of genetic composition in brown trout (*Salmo trutta* L.) populations inhabiting an unstable environment. Mol. Ecol. **12**: 3123-3135.

Poulsen, N.A., Nielsen, E.E., Schierup, M.H., Loeschcke, V., and Gronkjaer, P. 2006. Long-term stability and effective population size in North Sea and Baltic Sea cod (*Gadus morhua*) Mol. Ecol. **15**: 321-331.

Paetkau, D., Slade, R., Burdens, M., and Estoup, A. 2004. Genetic assignment methods for the direct, real time estimation of migration rate: a simulation-based exploration of accuracy and power. Mol. Ecol. **13**: 55–65.

Palm, S., Dannewitz, J., Jarvi, T., Petersson, E., Prestegaard, T., and Ryman, N. 2003. Lack of molecular genetic divergence between sea-ranched and wild sea trout (*Salmo trutta*). Mol. Ecol. **12**: 2057–2071.

Palstra, F.P., O'Connell, M.F., and Ruzzante, D.E. 2007 Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. Mol. Ecol. **16**: 4504-4522.

Palstra, F.P and Ruzzante, D.E. 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Mol. Ecol. **17**: 3428-3447.

Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes **6**: 288-295.

Peel, D., Ovenden, J.R., and Peel, S.L. 2004. Ne estimator: Software for Estimating Effective Population Size, Version 1.3. Queensland Government: Department of Primary Industries and Fisheries, Brisbane, Queensland.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., and Estoup, A. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. J. Hered. **95**: 536-539.

Quinn, T.P 1993. A review of homing and straying of wild and hatchery-produced salmon. Fish. Res. **18**: 29-44.

Rannala, B., and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. Proc. Natl. Acad. Sci. USA. **94**: 9197-201.

Raymond, M.L., and Rousset, F. 1995. An exact test for population differentiation. Evolution, **49**: 1280-1283.

Rhodes, K.L., Lewis, R.I., Chapman, R.W., and Sadovy, Y. 2004. Genetic structure of camouflage grouper, *Epinephelus polyphekadion* (Pisces: Serranidae), in the western central Pacific. Mar. Biol. **142**: 771-776.

Ruzzante, D.E., Taggart, C.T., Cook, D., and Goddard, S.E. 1997. Genetic differentiation between inshore and offshore Atlantic cod (*Gadus morhua*) off Newfoundland: a test and evidence of temporal stability. Can. J. Fish. Aquat. Sci. **54**: 2700–2708.

Schneider, S., Kueffer, J-M., Roesslo, D., and Excoffier, L. 2000. Arlequin, Version 2000. A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva.

Shrimpton, M.J., and Heath, D.D.2003. Census vs. effective population size in Chinook salmon: large- and small-scale environmental perturbation effects. Mol. Ecol. **12**: 2571-2583.

Slatkin, M. 1985. Gene flow in natural populations. Ann. Rev. Ecol .Syst. **16**: 393-430.

Taylor, E.B. 1991. A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture, **98**: 185-207

Taylor, E.B. 1995. Genetic variation at minisatellite DNA loci among north Pacific populations of steelhead and rainbow trout (*Oncorhynchus mykiss*). J. Hered. **86**: 354-363.

Tessier, N., and Bernatchez, L. 1999. Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (*Salmo salar* L.). Mol. Ecol. **8**: 169–179.

Wang, J., and Whitlock, M.C. 2003. Estimating effective population size and migration rates from genetic samples over space and time. Genetics, **163**: 429-446.

Waples, R.S. 1989. A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics, **121**: 379-391.

Weir, B.S., and Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution, **38**: 1358-1370.

Williams, C.L., Bleiwas, K, Johnston, J.J., and Jaeger, M.M. 2003. Temporal genetic variation in a coyote (*Canis latrans*) population experiencing high turnover. J. Mammology, **84**: 177-184.

Wilson, G.A., and Rannala, B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics, **163**: 1177–1191.

Table 2.1 Pairwise F_{ST} for spatiotemporal comparisons of five populations of Chinook salmon (*Oncorhynchus tshawytscha*) from the Upper Fraser River. n = sample sizes; all P-values for F_{ST} estimates are significant at $P < 0.01$, except bold values significant at $P < 0.05$ and underlined values are not significant $P > 0.05$.

	Nechako			**Dome**			**Willow**		**Bowron**			**Stuart**		
	N-78	**N-88**	**N-98**	**D-86**	**D-91**	**D-96**	**W-80**	**W-92**	**B-80**	**B-88**	**B-98**	**S-80**	**S-88**	**S-98**
n	49	60	48	29	58	59	27	30	36	34	18	55	44	43
N-78														
N-88	0.006													
N-98	0.009	0.014												
D-86	0.040	0.038	0.059											
D-91	0.025	0.025	0.034	0.017										
D-96	0.042	0.029	0.047	0.032	0.031									
W-80	0.018	0.021	0.024	0.040	0.038	0.036								
W-92	0.051	0.044	0.037	0.075	0.057	0.032	0.032							
B-80	0.048	0.052	0.058	0.056	0.051	0.041	0.032	0.040						
B-88	0.050	0.052	0.069	0.040	0.039	0.051	0.034	0.076	0.020					
B-98	0.032	0.033	0.043	0.041	**0.023**	0.046	**0.028**	0.053	0.044	0.025				
S-80	0.027	0.026	0.026	0.065	0.046	0.033	0.025	0.021	0.026	0.046	0.028			
S-88	0.029	0.024	0.024	0.069	0.046	0.042	0.040	0.030	0.047	0.066	0.045	**0.014**		
S-98	0.033	0.031	0.032	0.087	0.058	0.046	0.049	0.040	0.048	0.071	0.054	0.022	0.024	

Table 2.2 Estimates of migration rates among populations and sampling dates using **BAYESASS 1.3**. Bold values along the diagonal are the proportion of fish that were assigned to the site of capture and are thus non-migrant fish. Site codes are: N = Nechako, W = Willow, B = Bowron, S = Stuart, D = Dome, with years; 95- and 95+ correspond to 95% confidence limits; *m* = total migration rate into each population.

	Into this site:											
Source	**N-78**	95-	95+	**W-80**	95-	95+	**B-80**	95-	95+	**S-80**	95-	95+
N-78	**0.98667**	0.95996	0.99936	0.00498	0.00002	0.02667	0.00571	0.00004	0.02531	0.12467	0.07421	0.17297
W-80	0.00320	0.00001	0.01523	**0.97692**	0.92000	0.99918	0.09765	0.00309	0.19019	0.00404	0.00003	0.01744
B-80	0.00333	0.00001	0.01653	0.01026	0.00002	0.05605	**0.88204**	0.79293	0.98085	0.00582	0.00005	0.02283
S-80	0.00680	0.00002	0.02814	0.00783	0.00002	0.03767	0.01460	0.00010	0.05931	**0.86547**	0.81647	0.91653
m	0.01333			0.02308			0.11796			0.13453		

	Into this site:											
Source	**N-88**	95-	95+	**D-86**	95-	95+	**B-88**	95-	95+	**S-88**	95-	95+
N-88	**0.94075**	0.89629	0.97518	0.00495	0.00001	0.02205	0.00811	0.00006	0.03655	0.01772	0.00007	0.06096
D-86	0.04569	0.01864	0.08321	**0.67707**	0.66702	0.71099	0.11796	0.05783	0.22797	0.00561	0.00002	0.02226
B-88	0.00393	0.00003	0.01681	0.31200	0.26099	0.32890	**0.86736**	0.75720	0.92552	0.00841	0.00003	0.03710
S-88	0.00963	0.00006	0.03567	0.00456	0.00001	0.02138	0.00658	0.00005	0.03191	**0.96827**	0.91673	0.99855
m	0.05925			0.32151			0.13264			0.03173		

	Into this site:											
Source	**N-98**	95-	95+	**D-96**	95-	95+	**B-98**	95-	95+	**S-98**	95-	95+
N-98	**0.97704**	0.93524	0.99892	0.00261	0.00001	0.01369	0.02060	0.00008	0.07173	0.01729	0.00007	0.05531
D-96	0.00588	0.00003	0.02642	**0.98977**	0.96658	0.99943	0.01231	0.00006	0.05270	0.00529	0.00001	0.02513
B-98	0.00987	0.00005	0.03472	0.00290	0.00001	0.01494	**0.95681**	0.88818	0.99761	0.00399	0.00001	0.01835
S-98	0.00721	0.00002	0.03490	0.00472	0.00003	0.02106	0.01027	0.00004	0.04566	**0.97344**	0.93224	0.99858
m	0.02296			0.01023			0.04319			0.02656		

Table 2.3 Number of genotyped fish following removal of first-generation (F_0) migrants identified in each population by **GENECLASS**. Population names follow as listed in the text; n = total number of genotyped fish per population; Mean D_{LR} measures genetic distance between focal population and all others.

Population	n	F_0 Migrants	% Migrants	Mean D_{LR}	n without migrants
N-78	49	5	0.10	2.1	44
N-88	60	3	0.05	2.5	57
N-98	48	8	0.17	3.9	40
D-86	29	5	0.17	4.5	24
D-91	58	6	0.10	4.3	52
D-96	59	11	0.19	4.8	48
W-80	27	7	0.26	2.7	20
W-92	30	6	0.20	2.9	24
B-80	36	7	0.19	2.7	29
B-88	34	5	0.15	3.2	29
B-98	18	3	0.17	4.9	15
S-80	55	2	0.04	2.4	53
S-88	44	2	0.05	4.2	42
S-98	43	6	0.14	3.9	37

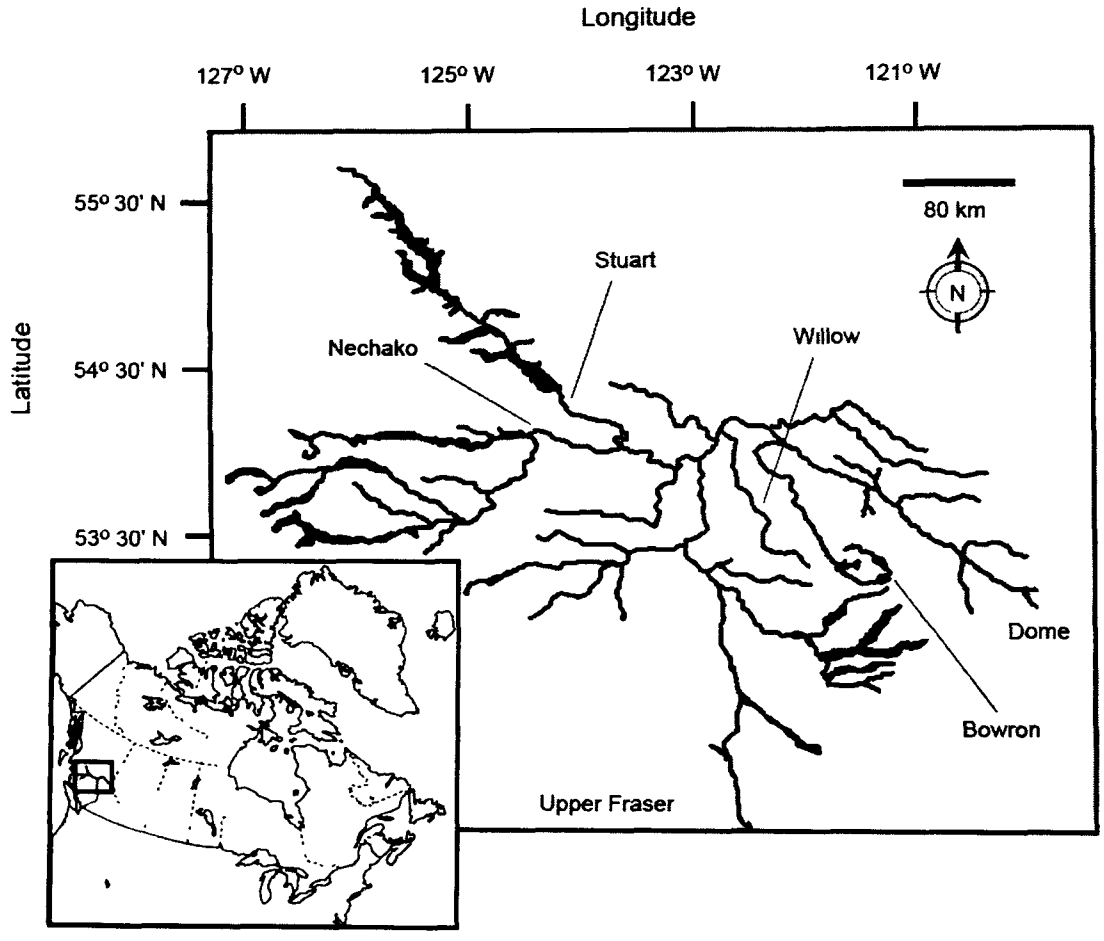


Figure 2.1 — Map of the upper Fraser River watershed showing the locations of the five rivers where spawning Chinook salmon samples were collected. Map redrawn from Shrimpton and Heath (2003).The Upper Fraser is marked in gray.

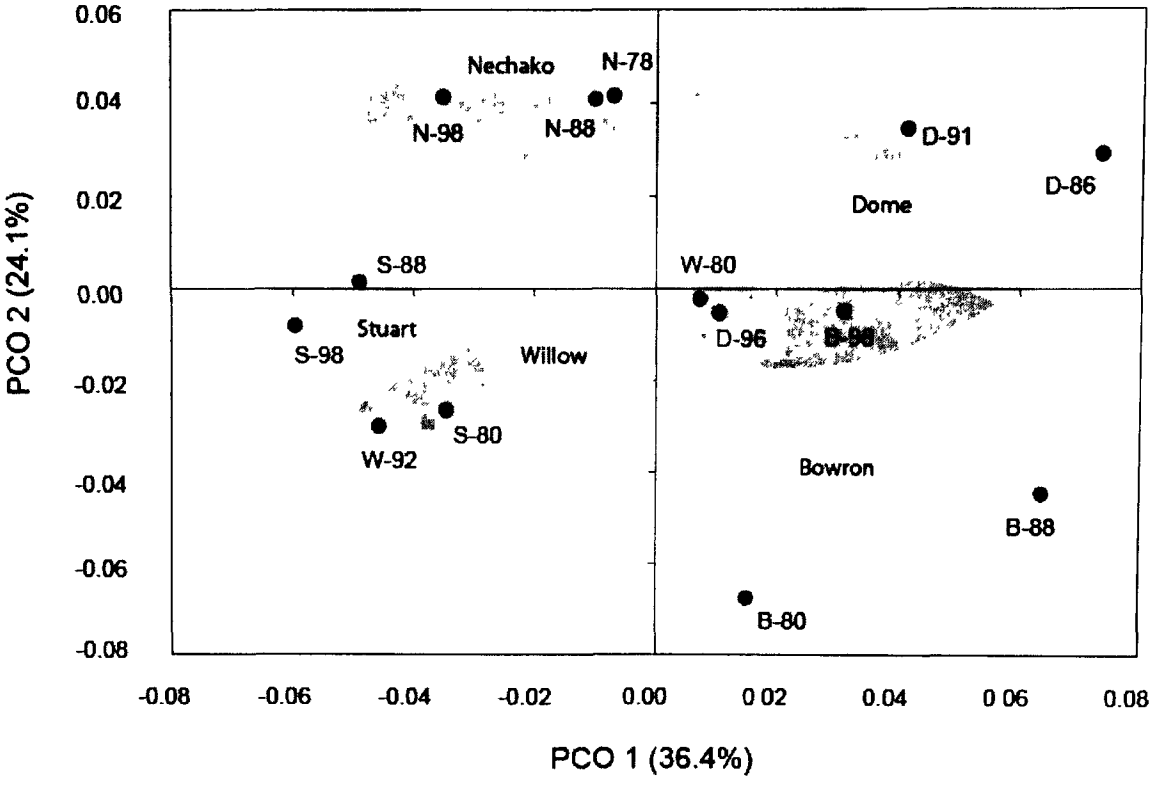


Figure 2.2 — Multidimensional scaling plot of pairwise F_{ST} among populations and sampling dates. Shaded ovals indicate spatial relationships among samples.

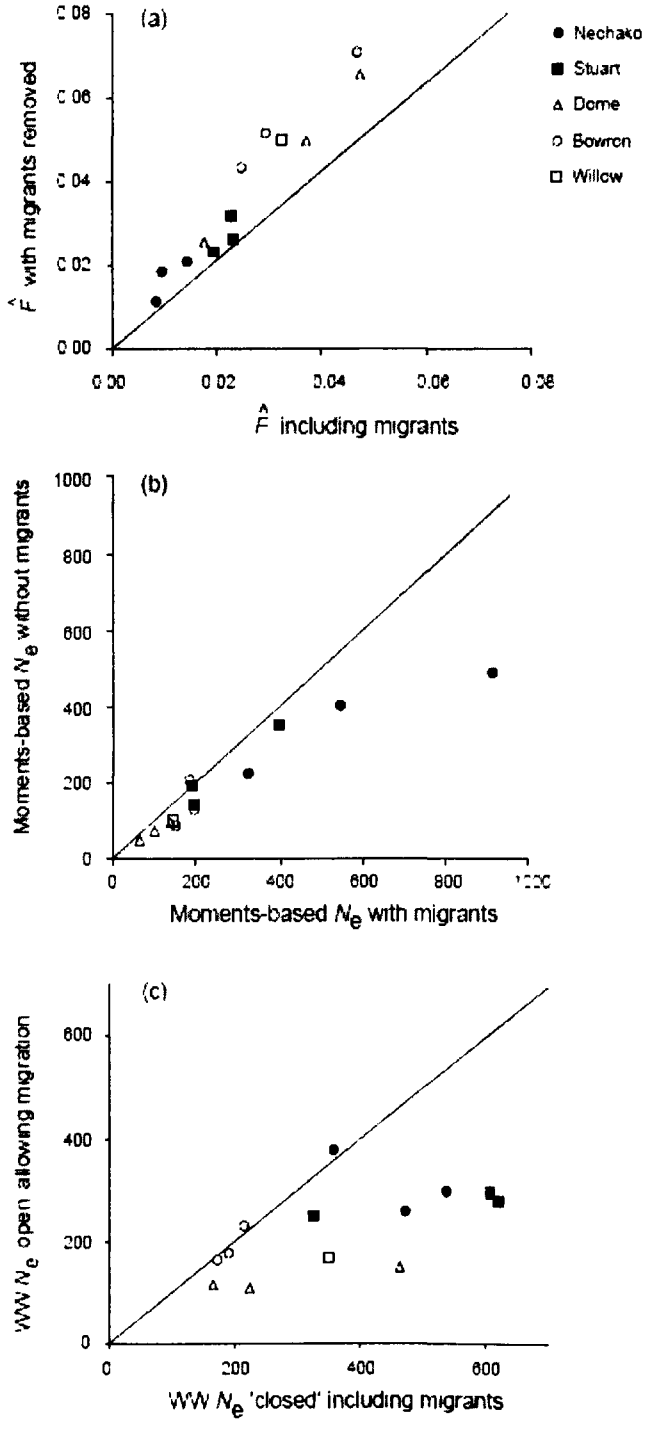


Figure 2.3 — Effects of migration on (a) time-standardized allelic variance ($\hat{F}$). (b) Translation of $\hat{F}$ estimates to Moments-based effective population size (N_e). (c) Comparison of Wang and Whitlock (WW, 2003) 'closed' estimates including migrants to WW 'open' model allowing migration. Straight lines indicate a 1:1 relationship.

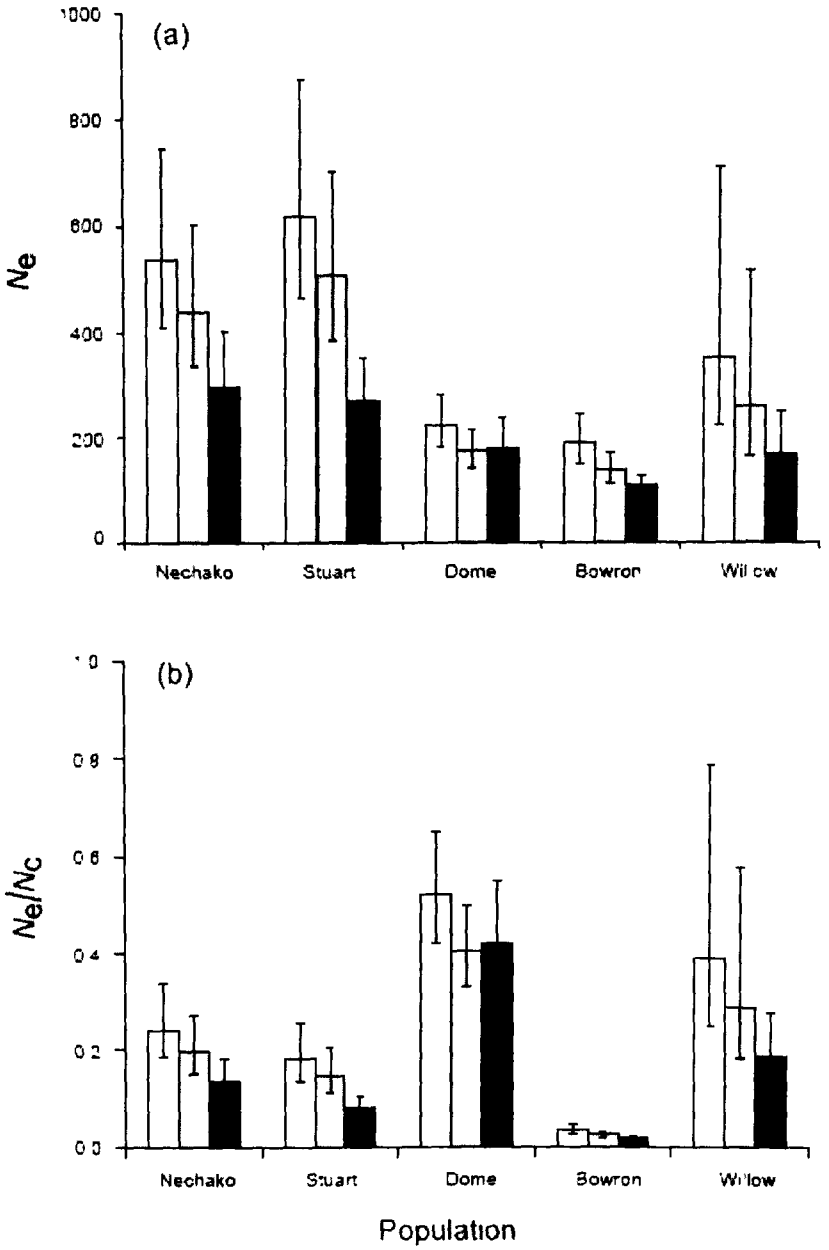


Figure 2.4 — (a) Study-wide comparative estimates of N_e using Wang and Whitlock (WW, 2003) 'closed' models both with migrants (white bars) and without migrants (light gray bars) and WW 'open' model allowing migration (dark gray bars). (b) Study-wide ratio of effective population size to census population size (N_e/N_c) estimates for N_e estimated assuming no migration and after correction for migration. Error bars represent 95% confidence limits.

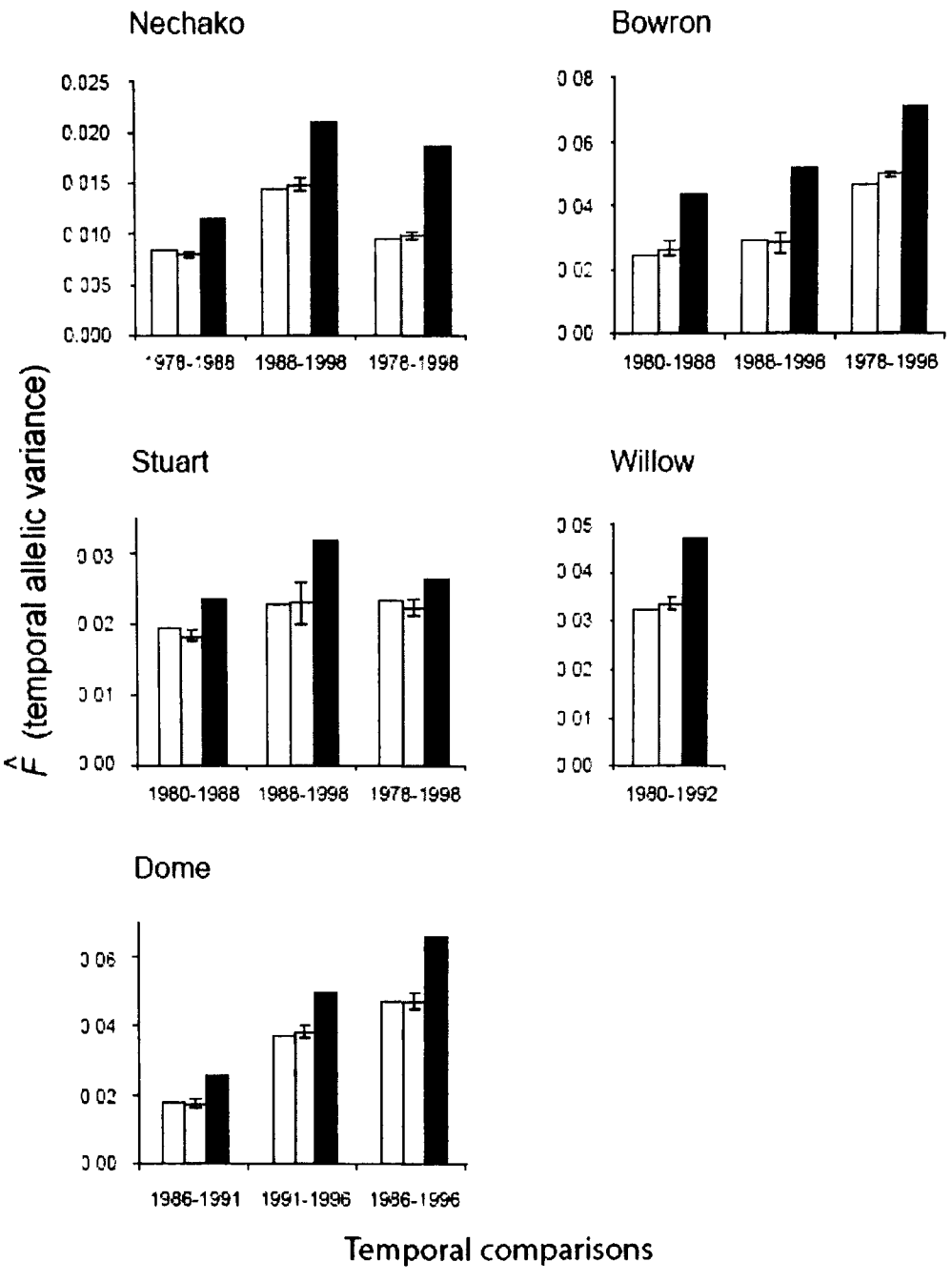


Figure 2.5 — Comparison of $\hat{F}$ estimates with migrants (white bars), removal of random individuals in equal proportion to detected migrants (gray bars; 10 replicates, error-bars represent SD), and removal of migrants detected by **GENECLASS 2.0** (black bars).

3.0 — DISPERSAL AND POPULATION GENETIC STRUCTURE OF 'SMALL' *TELMATHERINA ANTONIAE*, AN ENDEMIC FRESHWATER SAILFIN SILVERSIDE FROM SULAWESI, INDONESIA*

3.1 INTRODUCTION

Dispersal, the movement of individuals or their gametes from their site of origin, is a principal factor in the maintenance of genetic connectivity among populations through gene flow. Constraints on dispersal restrict gene flow resulting in the isolation of groups of individuals, thereby initiating population divergence and ultimately driving allopatric speciation (Bossart & Prowell, 1998; Bohonak, 1999).

Limits to dispersal are the result of extrinsic (physical barriers) or intrinsic (behavioural and/or ecological) factors. In general, genetic differentiation is more pronounced among conspecific populations of freshwater compared to marine species (Bilton *et al* , 2002; Wong *et al.*, 2004), and is largely attributed to geographic barriers (Ward & Elliott, 2001). Among lake-dwelling fishes population-level genetic structure ranges from lake-wide panmixia to highly divergent local populations depending on species-specific biology and lake-specific limnological characteristics (e.g. lake size, bathymetry and productivity). Within-species population structuring in a single lake has been documented for a number of fish species including cichlids (Shaw *et al* , 2000; Abila *et al.*, 2004), sardines (Hauser *et al.*, 1998), walleye (Stepian & Faber, 1998; DuPont *et al* , 2007), arctic charr and other salmonids (Power *et al.*, 2005; Hendry *et al.*, 1998, 1999). In those studies, spawning site fidelity, habitat partitioning, and microallopatry

* modified from Walter RP, Haffner GD, Heath DD 2008 Dispersal and population genetic structure of *Telmatherina antoniae*, an endemic freshwater Sailfin silverside from Sulawesi, Indonesia doi 10 1111/j 1420-9101 2008 01645.x

have been postulated as driving population structure.

In a number of African cichlid species, habitat-fidelity and microallopatry appear to be important predictors of population genetic structure. Habitat barriers including deep water, sandy areas and even habitat discontinuity have been identified as barriers to gene flow (Danley *et al.*, 2000; Rico & Turner, 2002). In the African Lake Tanganyika, high cichlid population differentiation was noted across barriers as well as along stretches of continuous shoreline (Taylor *et al.*, 2001; Rüber *et al.*, 2001; Duftner *et al.*, 2006). Given this influence of restricted dispersal and microallopatry in the strong population genetic divergence and speciation of African cichlids, similar patterns may exist for other adaptively radiating freshwater species flocks.

The telmatherinid species complex of the Malili Lakes, Sulawesi, Indonesia, have gained recent attention through evidence of ecological segregation (Roy *et al.*, 2007a, 2007b), hybridization (Herder *et al.*, 2006) and adaptive radiation (Roy *et al.*, 2004, Herder *et al.*, 2006); and have been compared to African cichlids as models for the study of adaptive divergence and speciation (Roy *et al.*, 2007a; 2007b; Herder *et al.*, 2006). Within the species flock are forms confined to a single lake, including *Telmatherina antoniae* Kottelat, endemic to Lake Matano (Kottelat *et al.*, 1993). Little is known about population genetic structure in the Lake Matano *telmatherinids*; however, the conspicuous littoral distribution of all seven described species indicates that the deep waters toward the center of the lake may serve as a barrier to dispersal, while other, perhaps less obvious barriers may also exist. Thus, philopatric behaviour of individuals may result in marked genetic divergence (and ultimately speciation) similar to that

noted in rock-dwelling cichlids (Taylor *et al.*, 2001; Rüber *et al.*, 2001; Rico & Turner, 2002; Sefc *et al.*, 2007; Duftner *et al.*, 2006).

Here, I quantify population structure and dispersal in *T. antoniae* using multilocus genotyping at polymorphic microsatellite loci coupled with a systematic sampling of the known species range (i.e., Lake Matano). Adult *T antoniae* form loose aggregations of 10-100 individuals and exhibit a patchy distribution as conspicuous members of Lake Matano's littoral fish community. However, aggregations do not appear to be associated with a distinct microhabitat. Furthermore, there is a lack of continuous littoral habitat throughout Lake Matano with some sites separated by deep drop-offs greater than 100 m (Fig. 3.1). Two questions form the focus of this study: 1) Given the presumed littoral distribution of this species is there genetic structure and/or dispersal among Lake Matano's *T. antoniae* populations? 2) Is there cryptic genetic structuring in the Lake Matano *T. antoniae* that does not conform to the spatial distribution of the fish in the littoral zone? As *T. antoniae* are endemic to Lake Matano, and are part of a species radiation, characterization of the factors shaping population genetic structure may provide important insights into mechanisms of sympatric speciation and local adaptation.

3.2 Materials and Methods

Collection and genetic analyses

Lake Matano, Sulawesi, Indonesia, is the hydrological head of the Malili Lakes and is characterized by a narrow and steep littoral zone (Roy, 2006; Fig. 3.1). Telmatherinids were captured by beach seine at depths ranging from 0-3 m from ten littoral sites in Lake Matano (Fig. 3.1) in June 2006. Adult 'torpedo' morph individuals with round-shaped 2nd dorsal and anal fins

were identified in the field as *T. antoniae* (following Kottelat, 1991, 1993; Roy 2006) and were sampled non-lethally by removing a small section of the anal fin tissue, which was preserved in 95% ethanol. The fish were released after a short recuperation period. DNA was recovered from tissue samples following the plate-based extraction method (Elphinstone *et al.*, 2003) and resuspended in 50 µl of Tris-EDTA buffer (10mM Tris, 1.0mM EDTA, pH 8.0).

To confirm our visual identification of *T. antoniae*, I genetically identified all fish based on mtDNA haplotype (Roy *et al.*, 2007a). I amplified a 1200 bp fragment of the cytochrome *b* gene using the forward primer GLUDG-5 (5'-TGACTTGAARAACCACCGTTG-3') (Palumbi, 1992) and reverse primer CBtelm-R (5'-GTGGAGGAGGGGTACGACTA-3') (Roy *et al.*, 2007a) using the PCR conditions: initial denaturation at 95°C for 2 min, followed by 35 cycles of 95°C for 10 s, 60°C for 10 s, 72°C for 10 s, followed by a final extension at 72°C for 1 min. Amplified *cyt b* fragments were then digested using PleI and FokI (New England Biolabs) to produce RFLP patterns diagnostic to *Telmatherina* mtDNA clades (following Roy, 2006). As introgressive hybridization has been documented in the adaptive radiation of Telmatherinid species complex (Herder *et al.*, 2006), any fish that did not conform to Roy's (2006) RFLP diagnostics were excluded from further analysis. Samples confirmed as *T antoniae* were then genotyped at eight polymorphic microsatellite loci (Tan9, Tan10, Tan11, Tan12, Tan14, Tan17, Tan24, Tan26) following Walter *et al.* (2007) using a LiCOR 4300 DNA analyzer and scoring allele size using GENE ImagIR 4.05 (Scanalytics Inc., Rockville, MD, USA) imaging software. I used MSA4.0 (Dieringer & Schlötterer, 2003) to look for genotyping errors, MICROCHECKER (van Oosterhout 2004) to check for the presence of null alleles, and GENEPOP4.0 (Raymond & Rousset, 1995) test for linkage disequilibrium among loci.

Data analyses

Non-equilibrium and equilibrium-based tests for genetic differentiation were performed. Pairwise Fisher's Exact tests (10000 dememorizations, 20000 permutations) were performed to test for significant differences in allele frequency distributions among sampled sites using TFPGA (Tools For Population Genetic Analysis, 1.3; Miller, 1997; Raymond & Rousset, 1995). Exact tests for HWE were also performed (20000 permutations) in TFPGA. Pairwise F_{ST} (Weir & Cockerham, 1984) was calculated using MSA4.0 (Dieringer & Schlötterer, 2003) and R_{ST} was calculated in ARLEQUIN 3.0 (Excoffier *et al.*, 2005). Significance values for Fisher's exact tests, HWE, and pairwise F_{ST} calculations were Bonferroni corrected.

Population genetic structure was tested for adherence to an isolation-by-distance model via two methods: 1) 'littoral restriction' (assumes fish avoid open water, with geographic distances calculated along the shortest littoral route between sites; and 2) 'open-water dispersal' using the shortest water distances between sites. Significance for IBD relationships were determined using Mantel tests (Mantel, 1967) in TFPGA.

Partial Bayesian genotype assignments (Rannala & Mountain, 1997) using population exclusion methods in GENECLASS 2.0 (Piry *et al.*, 2004) were performed to detect first generation migrants among sites using the likelihood ratio *L-home / L-max* (Paetkau *et al.*, 2004). Monte-Carlo resampling using Paetkau *et al.*'s (2004) simulation algorithm was performed with 10,000 simulated individuals at an assignment threshold P-value of 0.05. Individuals identified as migrants were assigned to a site of origin on the basis of their -ln likelihood. Mean dispersal

distances were calculated from the shortest straight-line distances between sampled sites where migrants originated and sites at which they were caught. Individual dispersal distances were compared to expected pairwise distances between all sites using a Kolmogorov-Smirnov test. I also used the migrant data to assess patterns in dispersal direction by assigning migrants back to their putative sources, then mapped the shortest-route distance onto a quadrant-map of Lake Matano to determine whether fish dispersed with respect to north, south, east or west direction (roughly along the long- and short-axes of the lake).

The Bayesian clustering programs STRUCTURE and BAPS work well for inferring the number of genetic clusters even at low levels of differentiation (F_{ST} = 0.02 to 0.03, Latch *et al.*, 2006). Lake-level population genetic structure was assessed independent of the spatial sampling pattern employed using *a posteriori* Bayesian genotype clustering in STRUCTURE 2.1 (Pritchard *et al.*, 2000). The analyses were performed in three independent runs using the admixture model with 500,000 burn-in followed by 1,000,000 MCMC repetitions, with three iterations (K ranging from 1 to 10). K was determined using ΔK following Evanno *et al.*, (2005). I also used BAPS 4.14 (Corander *et al.*, 2006) to independently estimate the number of genetic clusters. Initial clustering of individuals was performed in three replicates for each possible K (1-10). I then performed an admixture analysis where the minimum population size prior to estimating admixture was 10, in 50 iterations, with admixture simulated from 200 reference individuals in 50 iterations. Hierarchical AMOVA was performed in ARLEQUIN 3.0 to determine the proportion of genetic variance explained for both the STRUCTURE and BAPS groups, excluding individuals who failed assignment ($Q < 0.5$ to any cluster), versus the spatially-defined groups.

To determine the potential effects of migrants on population structure, the STRUCTURE, BAPS and IBD analyses were re-run with migrant fish (identified as above) removed from datasets.

3.3 RESULTS

A total of 486 *Telmatherina antoniae* were genotyped at eight microsatellite loci with 3 to 25 alleles per locus and observed (H_O) and expected heterozygosities (H_E) ranging from 0.30 to 1.0 (Table 3.1). Exact tests revealed significant differences in allele frequency distributions among all pairs of sample sites after Bonferroni correction ($P < 0.001$). Ten of 90 tests revealed significant departures from HWE following Bonferroni correction, but none were consistent across sites or loci (Table 3.1). Pairwise F_{ST} estimates ranged from 0.009 - 0.061 (Table 3.2) with an overall significant Global F_{ST} estimate of 0.029 ($P = 0.0001$). Pairwise comparisons that included the Kupu site produced the highest F_{ST} estimates (Table 3.2, Fig 3.1). R_{ST} were 0.43% higher on average across all sites, and were comparable to F_{ST} estimates (Table 3.2).

Mantel tests for IBD relationships using genetic distance ($F_{ST}/(1-F_{ST})$) versus geographic distances (or log geographic distances) did not yield significant results under scenarios of straight-line dispersal ($P = 0.231$), nor littoral restriction dispersal ($P = 0.181$). There was no difference in the results of IBD analyses following removal of migrant fish from the dataset.

Migrant analysis and dispersal estimates

Detection of first-generation migrants using GENECLASS resulted in 50 individuals being classified as migrants, with 3-9 migrants per site (Table 3.3). Estimated dispersal distances ranged from 3.5 to 28.8 km with a mean estimated dispersal distance of 13.6 (SD = 0.94) km

(Table 3.3). Dispersal distances for *T. antoniae* were not significantly different from random expected open-water distances using a Kolmogorov-Smirnov test ($d_{max\ 0\ 05,7,45} = 6$, $P > 0.05$, Fig 3.2). No significant directional bias relative to random expectations was observed.

Lake-level genetic structuring

Individual-based genotype clustering using STRUCTURE returned similar –ln likelihood values for $K = 4$, $K = 5$, and $K = 6$; however, those clusters did not correspond to any of the sample sites or geographic structuring. I selected $K = 4$ as the estimate of the correct number of genetic clusters using both the ΔK criterion of Evanno *et al.* (2005) and the recommendations of Pritchard *et al.* (2000). I then determined the genetic cluster composition at each sample site, following removal of 65 individuals (13% study wide) identified as unassigned (e.g., these include "admixed" individuals) by a $Q < 0.5$ to any given cluster. I also forced assignment of all individuals on the basis of their highest Q coefficient for a cluster. No gender- or size-based differences were found between clusters. Clustering using BAPS also failed to produce genetic structure corresponding to geographic structuring; with $K = 6$, of which the two additional groups primarily nested within STRUCTURE's groups 1 and 4. Both Bayesian programs grouped the same individuals into group 2, indicating high divergence of these individuals with respect to others in the dataset. BAPS identified 11 individuals as significantly admixed (thus unassigned) using BAPS at $P < 0.05$; however if the threshold is raised to $P < 0.10$, the number of significantly admixed individuals increases to 52, closer to the number identified by the STRUCTURE analysis. Significant differences in the proportion of each population cluster were apparent at each sampling site ($\chi^2_{27} = 134$, $P < 0.0001$) using both Bayesian clustering programs (Fig. 3.3, only STRUCTURE results shown).

From the hierarchical AMOVA, the genetic variance explained by the Bayesian clusters was two times greater using BAPS than geographic groupings (5.42 %, $P < 0.001$; 2.32%, $P < 0.001$, respectively) and three times greater using STRUCTURE (7.30%, $P < 0.001$; 2.26%, $P < 0.001$, respectively). F_{ST} estimates between clusters were also larger (Table 3.4) in comparison to the geographic estimates (Table 3.2) with cluster 2 (both STRUCTURE and BAPS) as the most divergent. I reran the AMOVA using the data including forced assignment of individuals; the results were nearly identical with 7.30% of the variance explained by STRUCTURE's clusters, with only minor differences in spatial variance (2.92%).

3.4 DISCUSSION

Among freshwater fish species, the degree of population sub-structure within a lake varies considerably. The majority of studies of within-lake genetic structure involving non-temperate species have focused on the cichlid species complexes of the African Rift lakes, with some species exhibiting little to no lake-wide structuring (Taylor & Verheyen, 2001; Koblmüller *et al.*, 2007) and others showing substantial genetic structure on small spatial scales (Taylor *et al.*, 2001; Rico & Turner, 2002; Duftner *et al.*, 2006; Sefc *et al.*, 2007). In the present study, I found significant spatial genetic structure in *T. antoniae* within a single lake, however; results of my migrant analysis and cluster-based assignments suggests that deep water and/or habitat discontinuity do not represent barriers to dispersal. Adjacent sites lacking a continuous littoral habitat and separated by water exceeding 100 m in depth were among the least spatially differentiated (e.g., Tanah Merah vs. Owesu: $F_{ST} = 0.01$; Table 3.1). Mean estimated dispersal distances spanned approximately half the length of the lake, indicating a lack of strict allopatry

and the opportunity for considerable gene flow. Furthermore, the Lake Matano *T. antoniae* do not conform to an isolation-by-distance model of population structure, making it unlikely that the observed genetic structure is due to distance-related reproductive isolation. I therefore conclude that although constraints to gene flow exist for *T. antoniae,* on-going dispersal is occurring throughout Lake Matano.

The dispersal capacity of *T. antoniae* may be characteristic of Atheriniformes in general and the retention of such a trait would be particularly favoured in habitats where the cost of dispersal was low. Dispersal behaviour is risky given possibilities of failure to locate suitable habitat and mates, or *en route* mortality. The low productivity of Lake Matano (Haffner *et al.,* 2001) is characterized by an absence of pelagic predators, which would foster movement among sites thereby minimizing the predation risk of dispersal. However, it appears that some intrinsic factor(s) maintain significant genetic structuring despite such dispersal.

Local adaptation could account for the low but significant spatial structuring, as dispersing individuals (e.g., immigrants) may exhibit reduced fitness compared to residents (Hendry *et al.,* 2002). However, the low number of identified migrant individuals in this study is likely the result of reduced power in the assignment analyses because of low genetic differentiation among sites. Therefore, the true number of dispersing individuals is likely higher and cannot be definitively estimated due to spatial overlap of more than one population.

Multilocus genetic assignment methods such as those used in this study provide opportunities for defining populations with uncertain physical boundaries (Manel *et al.,* 2005; Rowe &

Beebee, 2007). The genotype-based clustering explained a greater proportion of the genetic variance than the spatially-based groups, indicative that the genetic clusters are present at all sites, spanning discontinuous habitat. Thus the sampled fish from the ten sites consist of differential proportions of individuals from the divergent and dispersive genetic clusters; however, the degree of isolation between the genetic clusters is only roughly estimated here. This type of population structure is analogous to a 'mixed-stock fishery' where different mixtures of divergent populations (or "stocks") result in genetic structure among samples. This type of genetic structure is common for a number of marine and aquatic organisms exhibiting philopatry, notably highly migratory sea turtles (Bowen & Karl, 2007), but has also been noted in landlocked salmonids (Potvin & Bernatchez, 2001). However, this observed among-sample site genetic divergence is problematic as it violates population genetic assumptions as specified by the Hardy-Weinberg principle.

First, the overlap of two or more distinct populations within a sample leads to a heterozygote deficit potentially creating a Wahlund effect, which may explain some departures from HWE among the data (Johnson & Black, 1984). The estimation of K based on Evanno *et al.* (2005) and Pritchard *et al.* (2000), coupled with the assignment criteria of Latch *et al.* (2006) used here appears to be robust given concordance of STRUCTURE and BAPS results. Secondly, while the sympatry of these genetic clusters serves to explain the observed spatial genetic structure with high dispersal, it does not explain the persistence of these genetic clusters despite lake-wide dispersive ability. Thus, this apparent mixture of populations among sites may have resulted from two processes: 1) secondary contact following past allopatric divergence, or 2) on-going reproductive isolation and genetic divergence.

The first possibility is less likely on two grounds. First, Lake Matano is considered an ancient lake (~4.5 my) with no evidence to suggest that lake-level fluctuations have produced isolated refugia within Lake Matano at any time in its history Furthermore, the estimated divergence times between *T. antoniae* and its sympatric sister mtDNA clade is 0.95–1.9 mya (Roy *et al.*, 2007a). Thus there no evidence for past vicariance affecting Lake Matano *T antoniae* populations, but it is a possibility. These markers are based on rapidly evolving microsatellite DNA, therefore gene flow facilitated by high dispersal may rapidly erode past allopatric divergence. Secondly, Roy *et al.* (2007a) also showed that the three mitochondrial telmatherinid clades likely diverged from a common ancestor within Lake Matano on the basis of ecological segregation, suggesting that any subsequent genetic structuring within *T. antoniae* may have occurred within Lake Matano as a result of similar processes.

The presence of sympatric but divergent genetic clusters may be the result of a number of intrinsic isolating mechanisms. For example, individuals may assortatively mate according to phenotype (e.g. colour, behavioural displays), as sexual selection has been implicated as an important factors driving sympatric genetic differentiation in other fishes (Seehausen, 2000). For *T. antoniae,* courtship and nuptial behaviours appear quite complex, with competition noted among males for access to females (Gray & McKinnon, 2006). Roy *et al.* (2007b) argued that colour was not an initiating factor in the earlier telmatherinid divergence, but did not assess colour as a potential structuring factor within specific telmatherinid clades. More recently, Herder *et al.* (2008) demonstrated a lack of colour-specific population differentiation among "small" *T. antoniae* colour morphs using AFLP markers. However, for the closely-related

Telmatherina sarasinorum Gray *et al.* (2008) provide evidence for environment-contingent sexual selection favouring different colour morphs in various habitats, and it is possible that a similar pattern is operating for *T antoniae*.

Herder *et al.* (2008) has further provided evidence for significant genetic differentiation among 'large' and 'small' morphtypes of *T. antoniae* and a lack of significant divergence between 'small' *T. antoniae* and the rare closely related but larger-sized *Telmatherina prognatha*. It is not without possibility that immature 'large' *T. antoniae* and *T prognatha* morphs may account for some of the genetic structure observed here. However, within our samples, spatially- or temporally-isolated reproductive aggregations could produce the observed genetic structuring. Genetic divergence at neutral microsatellite loci among populations of marine Atheriniformes with high dispersal capacity has classically been interpreted as a result of homing behaviour (Beheregaray & Sunnucks, 2001). However, little is known of the reproductive biology of *T. antoniae*, and the apparent lack of habitat-specificity precludes the identification of natal sites for this species.

The pattern of 'divergence despite gene dispersal' observed in this study has also been identified in incipient species complexes where strong divergent selection drives isolation despite on-going gene flow (Rice & Hostert, 1993). Seemingly sympatric populations may in fact be isolated by extreme philopatry or microgeographic structuring. This does not appear to be the case for *T. antoniae* where the lack of physical barriers to dispersal and the persistence of substantial genetic divergence lend support to the role of intrinsic mechanisms in the ongoing divergence among lineages in this ancient tropical lake.

3.5 REFERENCES

Abila R., Barluenga, M., Engelken, J., Meyer, A., & Salzburger, W. 2004 Population-structure and genetic diversity in a haplochromine fish cichlid of a satellite lake of Lake Victoria. *Mol. Ecol.* **13**: 2589-2602.

Bilton, D.T., Paula, J. & Bishop, J.D.D. 2002. Dispersal, genetic differentiation and speciation in estuarine organisms. *Estuar. Coast. Shelf. S.* **55**: 937–952.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. *Q. Rev. Biol.* **74**: 2145.

Bossart, J.L. & Prowell, D.P. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends Ecol. Evol.* **13**: 202-206.

Beheregaray, L.B. & Sunnucks, P 2001. Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. *Mol. Ecol.* **10**: 2849-2866.

Bowen, B.W. & Karl, S.A. 2007. Population genetics and phylogeography of sea turtles. *Mol. Ecol.* **16**: 4886-4907.

Corander, J. and Marttinen, P. 2006. Bayesian identification of admixture events using multi-locus molecular markers. *Mol. Ecol.* **15**: 2833-2843.

Danley, P.D., Markert, J.A., Arnegard, M.E. & Kocher, T.D. 2000. Divergence with gene flow in rock-dwelling cichlids of Lake Malawi. *Evolution* **54**: 1725-1737.

Dieringer, D. & Schlötterer, C. 2003. Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* **3**: 167-169.

Duftner, N., Sefc, K.M., Koblmüller, S., Nevado, B., Verheyen, E., Phiri, H. & Sturmbauer, C. 2006. Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. *Mol. Ecol.* **15**: 2381–2395.

DuPont, P-P., Bourret, V & Bernatchez, L. 2007. Interplay between ecological, behavioural and historical factors in shaping the genetic structure of sympatric walleye populations *Sander vitreus*. *Mol. Ecol.* **16**: 937-951.

Elphinstone, M.S., Hinten, G.N., Anderson, M.J. & Nock, C.J. 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Mol. Ecol. Notes*. **3**: 317-320.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Excoffier, L., Laval, G. & Schneider, S. 2005. Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.

Gray, S.M. & McKinnon, J.S. 2006. A comparative description of mating behaviour in the endemic Telmatherinid fishes of Sulawesi's Malili Lakes. *Environ. Biol. Fish.* **75**: 471-482.

Gray, S.M., Dill, L.M., Tantu, F Y., Loew, E.R., Herder, F & McKinnon, J.S. 2008. Environment contingent sexual selection in a colour polymorphic fish. *Proc. Roy. Soc. B* **275**: 1785-1791.

Haffner, G.D., Hehanussa, P.E. & Hartoto, D. 2001. The biology and physical properties of large lakes of Indonesia. In: *The Great Lakes of the World: Food-Web Health and Integrity* eds Manwar M, Heck RE, pp. 183-192. Backuys, Leiden.

Hauser, L., Carvalho, G.R. & Pitcher, T.J. 1998. Genetic population structure in the Lake Tanganyika sardine *Limnothrissa miodon. J. Fish. Biol.* **53**: Supplement A, 413-429.

Hendry, A.P., Hensleigh, J.E. & Reisenbichler, R.R. 1998. Incubation temperature, developmental biology and the divergence of sockeye salmon within Lake Washington. *Can. J. Fish. Aquat. Sci.* **55**: 1387-1394.

Hendry, A.P., Berg, O.K. & Quinn, T.P 1999. Condition dependence and adaptation-by-time: breeding date, life history, and energy allocation in a population of salmon. *Oikos* **85**: 499-514.

Hendry, A.P., Taylor, E.B. & McPhail, J.D. 2002. Adaptive divergence and the balance between selection and gene flow: lake and stream stickleback in the Misty system. *Evolution* **56**: 119-1216.

Hendry, A.P & Day, T 2005. Population structure attributable to reproductive time: isolation by time and adaptation by time. *Mol. Ecol.* **14**: 901–926.

Herder, F., Nolte, A.W., Pfaender, J., Schwarzer, J., Hadiaty, R.K. & Schliewen, U. 2006. Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc. Roy. Soc. B* **273**: 2209-2217

Herder, F., Pfaender, J. & Schliewen, U.K. 2008. Adaptive sympatric speciation of polychromatic "roundfin" sailfin silverside fish in Lake Matano. *Evolution* **62**: 2178-2195.

Koblmüller, S., Sefc, K.M., Duftner, N., Warum, M. & Sturmbauer, C. 2007. Genetic population structure as an indirect measure of dispersal ability in a Lake Tanganyika cichlid. *Genetica* **130**: 121-131.

Kottelat, M. 1991. Sailfin silversides Pisces: Telmatherinidae of Lake Matano, Sulawesi, Indonesia, with descriptions of six new species. *Ichthyol Explor Fres.* **1**: 321–344.

Kottelat, M., Whitten, A.J., Kartikasari, S.N. & Wirjoatmodjo, S. 1993. Freshwater Fishes of Western Indonesia and Sulawesi. Periplus, Jakarta.

Johnson, M.S. & Black, R. 1984. The Wahlund effect and the geographical scale of variation in the intertidal limpet *Siphonaria* sp. *Mar. Biol.* **79**: 295-302.

Latch, E.K., Dharmarajan, G., Glaubitz, J.C. & Rhodes, O.E. Jr. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conserv. Gen.* **7**: 295-302.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**: 209-220.

Manel, S., Gaggiotti, O.E. & Waples, R.S. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol.* **20**: 136–142.

Miller, M. 1997 Tools for population genetic analysis (TFPGA) 1.3: A Windows program for analysis of allozyme and molecular population genetic data. Computer software distributed by the author.

Paetkau, D., Slade, R., Burdens, M. & Estoup, A. 2004. Genetic assignment methods for the direct, real time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* **13**: 55–65.

Palumbi, S.R. 1992. Marine speciation on a small planet. *Trends Ecol. Evol.***7**: 114-118.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L., & Estoup A. 2004. GeneClass2: A software for genetic assignment and first-generation migrant detection. *J. Heredity* **95**: 563-539.

Potvin, C., & Bernatchez, L. 2001. Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multi-locus individual assignment and mixed-stock analyses. *Mol. Ecol.* **10**: 2375– 2388.

Power, M., O'Connell, M.F & Dempson, J.B. 2005. Ecological segregation within and among Arctic char morphotypes in Gander Lake, Newfoundland. *Environ. Biol. Fish.* **73**: 263-274.

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

Rannala, B. & Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. *P Natl. Acad. Sci-Biol.* **94**: 9197-9201.

Raymond, M. & Rousset, F 1995. Genepop Version 1.2: population genetics software for exact tests and ecumenicism. *J. Heredity* **86**: 248-249.

Rice, W.R. & Hostert, E.E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**: 1637-1653.

Rico, C. & Turner, G.F 2002. Extreme microallopatric divergence in a cichlid species from Lake Malawi. *Mol. Ecol.* **11**: 243-251.

Rowe, G. & Beebee, T.J.C. 2007. Defining population boundaries: use of three Bayesian approaches with microsatellite data from British natterjack toads *Bufo calamita*. *Mol. Ecol.* **16**: 785-796.

Roy, D., Docker, M.F., Hehanussa, P., Heath, D.D. & Haffner, G.D. 2004. Genetic and morphological data supporting the hypothesis of adaptive radiation in the endemic fish of Lake Matano. *J. Evol. Biol.* **17**: 1268-1276.

Roy, D. 2006. The evolutionary history and ecology of *Telmatherina* in Lake Matano: An example of adaptive radiation in an ancient lake. PhD Thesis, University of Windsor, Windsor, ON.

Roy, D., Paterson, G., Hamilton, P.B., Heath, D.D. & Haffner G.D. 2007a. Resource-based adaptive divergence in the freshwater fish *Telmatherina* from Lake Matano, Indonesia. *Mol. Ecol.* **16**: 35-48.

Roy, D., Docker, M.F., Haffner, G.D. & Heath, D.D. 2007b. Body shape vs. colour associated initial divergence in the *Telmatherina* radiation in Lake Matano, Sulawesi, Indonesia. *J. Evol. Biol.* **20**: 1126–1137.

Rüber, L., Meyer, A., Sturmbauer, C. & Verheyen, E. 2001. Population structure in two sympatric species of the Lake Tanganyika cichlid tribe Eretmodini: evidence for introgression. *Mol. Ecol.* **10**: 1207–1225.

Seehausen, O. 2000. Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.* **31**: 237-274.

Sefc, K.M., Baric, S., Salzburger, W & Sturmbauer, C. 2007. Species-specific population structure in rock-specialized cichlid species in Lake Tanganika, East Africa. *J. Mol. Evol.* **64**: 33-49.

Shaw, P W., Turner, G.F., Idid, M.R., Robinson, R.L. & Carvalho, G.R. 2000. Genetic population structure indicates sympatric speciation of Lake Malawi pelagic cichlids. *Proc. Roy. Soc. B* **267**: 2273–2280.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457-462.

Stepian, C.A. & Faber, J.E. 1998. Population genetic structure, phylogeography and spawning philopatry in walleye Stizostedion vitreum from mitochondrial DNA control region sequences. Mol. Ecol. 7: 1757-1769.

Taylor, M.I., Rüber, L. and Verheyen, R. 2001. Microsatellites reveal high levels of population substructuring in the species-poor Eretmodine cichlid lineage from Lake Tanganyika. *Proc. Roy. Soc. B* **268:** 803-808

Taylor, M.I. & Verheyen, E. 2001. Microsatellite data reveals weak population substructuring in *Copadichromis* sp. 'viriginalis kajose', a demersal cichlid from Lake Malawi, Africa. *J. Fish. Biol.* **59**: 593–604.

van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. 2004. MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**: 535–538.

Walter, R.P., Ovenden, J.R., & Heath, D.D. 2007. Characterization of four tetranucleotide and six dinucleotide microsatellite markers for use in *Telmatherina antoniae* and related species. *Mol. Ecol. Notes* **7**: 651-653.

Ward, R. D. & Elliot, N. G. 2001. Genetic population structure of species in the South East Fishery of Australia. *Mar. Freshwater Res.* **52**: 563–573.

Weir, B.S. & Cockerham, C.C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

Wong, B.B.M., Keogh, J.S. & McGlashan, D. J. 2004. Current and historical patterns of drainage connectivity in eastern Australia inferred from population genetic structuring in a widespread freshwater fish *Pseudomugil signifer* (Pseudomugilidae). *Mol. Ecol.* **13**: 391–401.

Table 3.1 Sample size (N), number of alleles (A), observed (H_O) and expected heterozygosities (H_E) for eight microsatellite loci in *Telmatherina antoniae* collected from 10 sites in Lake Matano, Indonesia. Departures from HWE (following Bonferroni correction) are underlined in bold.

		Sampling Site									
Locus		Lawa	P Rio	Taima	Indah	Kupu	L Ondau	Owesu	T Merah	W Lonto	Soluru
Tan 9	N	46	44	43	56	58	43	54	47	58	7
	A	24	20	22	21	25	24	22	19	24	8
	H_O	0.89	0.82	0.86	0.95	**0.88**	0.91	0.93	0.83	0.84	0.71
	H_E	0.93	0.94	0.93	0.94	**0.94**	0.95	0.92	0.93	0.92	0.89
Tan 10	N	49	52	47	57	53	46	56	50	56	8
	A	5	4	3	4	3	4	4	6	4	4
	H_O	**0.53**	0.40	0.36	0.49	0.47	0.30	0.55	0.36	0.39	0.38
	H_E	**0.58**	0.37	0.51	0.47	0.59	0.69	0.61	0.56	0.47	0.71
Tan 11	N	46	53	45	56	58	42	55	47	55	8
	A	10	9	10	11	16	10	10	10	14	6
	H_O	**0.41**	0.43	0.44	0.57	0.55	0.57	0.62	0.40	0.53	0.75
	H_E	**0.67**	0.63	0.70	0.67	0.77	0.7	0.71	0.69	0.81	0.73
Tan 12	N	50	53	47	57	58	45	57	49	58	8
	A	13	9	14	14	9	16	8	14	15	9
	H_O	**0.72**	0.64	0.72	0.54	0.55	**0.51**	**0.51**	0.73	0.60	0.75
	H_E	**0.83**	0.78	0.72	0.73	0.79	**0.82**	**0.8**	0.84	0.79	0.91
Tan 14	N	50	52	45	57	58	45	56	49	58	8
	A	8	7	7	7	7	5	7	5	7	4
	H_O	0.48	0.54	0.60	0.63	0.59	0.51	0.54	0.51	0.52	0.63
	H_E	0.74	0.68	0.76	0.70	0.73	0.69	0.73	0.63	0.72	0.69

Tan 17	*N*	51	50	46	57	58	45	57	50	59	8
	A	25	19	22	25	19	21	23	22	20	11
	H_O	0.9	0.92	0.91	0.86	0.98	0.89	0.84	0.88	0.9	1
	H_E	0.93	0.93	0.92	0.92	0.93	0.94	0.93	0.93	0.92	0.95
Tan 24	*N*	51	52	46	56	57	46	57	47	57	8
	A	8	8	8	9	7	9	7	7	8	7
	H_O	0.49	0.52	0.46	0.57	**0.4**	0.41	0.53	0.49	0.44	0.50
	H_E	0.57	0.53	0.67	0.60	**0.58**	0.61	0.49	0.55	0.56	0.79
Tan 26	*N*	51	53	47	57	58	46	57	49	58	8
	A	23	22	22	25	20	24	23	21	22	12
	H_O	0.96	0.89	0.94	0.95	1	0.89	0.89	0.88	0.97	1
	H_E	0.95	0.95	0.95	0.94	0.94	0.94	0.95	0.94	0.94	0.94

Table 3.2 Pairwise F_{ST} (below diagonal) and R_{ST} estimates for 10 sampled sites of *Telmatherina antoniae*. F_{ST} estimates significance levels < 0.05 are in bold, < 0.01 indicated by * following Bonferroni correction.

Site	Lawa	P Rio	Taima	Indah	Kupu	L Ondau	Owesu	T Merah	W Lonto	Soluru
Lawa		0.020	0.043	0.012	0.043	0.011	0.038	0.031	0.022	0.047
P Rio	**0.017***		0.039	0.022	0.055	0.037	0.031	0.029	0.027	0.086
Taima	**0.038***	**0.034***		0.040	0.054	0.035	0.015	0.017	0.030	0.072
Indah	0.010	**0.021***	**0.037***		0.053	0.017	0.047	0.032	0.022	0.051
Kupu	**0.043***	**0.051***	**0.050***	**0.050***		0.038	0.042	0.041	0.055	0.053
L Ondau	0.009	**0.034***	**0.030***	**0.015**	**0.036***		0.033	0.021	0.014	0.029
Owesu	**0.035***	**0.028***	**0.012**	**0.044***	**0.040***	**0.030***		0.014	0.027	0.083
T Merah	**0.029***	**0.025***	**0.014**	**0.031***	**0.039***	**0.018***	0.010		0.015	0.047
W Lonto	**0.021***	**0.024***	**0.026**	**0.019***	**0.051***	0.013	**0.025***	**0.013**		0.044
Soluru	0.037	**0.070***	**0.058**	**0.044***	**0.044**	0.018	**0.069***	0.038	0.036	

Table 3.3 Summary of migrant analysis (exclusion analysis) using GENECLASS 2.0, mean dispersal estimates taken from migrant individuals assigned back to their site of origin.

	Source Site										
	Lawa	P Rio	Taima	Indah	Kupu	L Ondau	Owesu	T Merah	W Lonto	Soluru	Mean distance (km)
Capture Site											
Lawa			1	1		1		1	1		18.4
P Rio			1	1			1		1		16.0
Taima	1			1	4				1		7.5
Indah			1		1	1	1				9.0
Kupu	1					1		2		1	10.5
L Ondau	3	2	1					2	1		17.9
Owesu	1		1							1	15.8
T Merah		1				1	1		1		10.6
W Lonto	1		2	1	1	1					15.1
Soluru				1				1	1		13.0

Table 3.4 Mean and pairwise F_{ST} estimates for Bayesian clusters generated from STRUCTURE ($K = 4$) and BAPS ($K = 6$), all estimates are highly significant following Bonferonni correction ($P < 0.001$).

STRUCTURE $F_{ST} = 0.102$

	1	2	3
2	0.120		
3	0.060	0.144	
4	0.031	0.181	0.079

BAPS $F_{ST} = 0.087$

	1	2	3	4	5
2	0.157				
3	0.034	0.103			
4	0.111	0.158	0.083		
5	0.072	0.146	0.047	0.047	
6	0.011	0.152	0.019	0.107	0.070

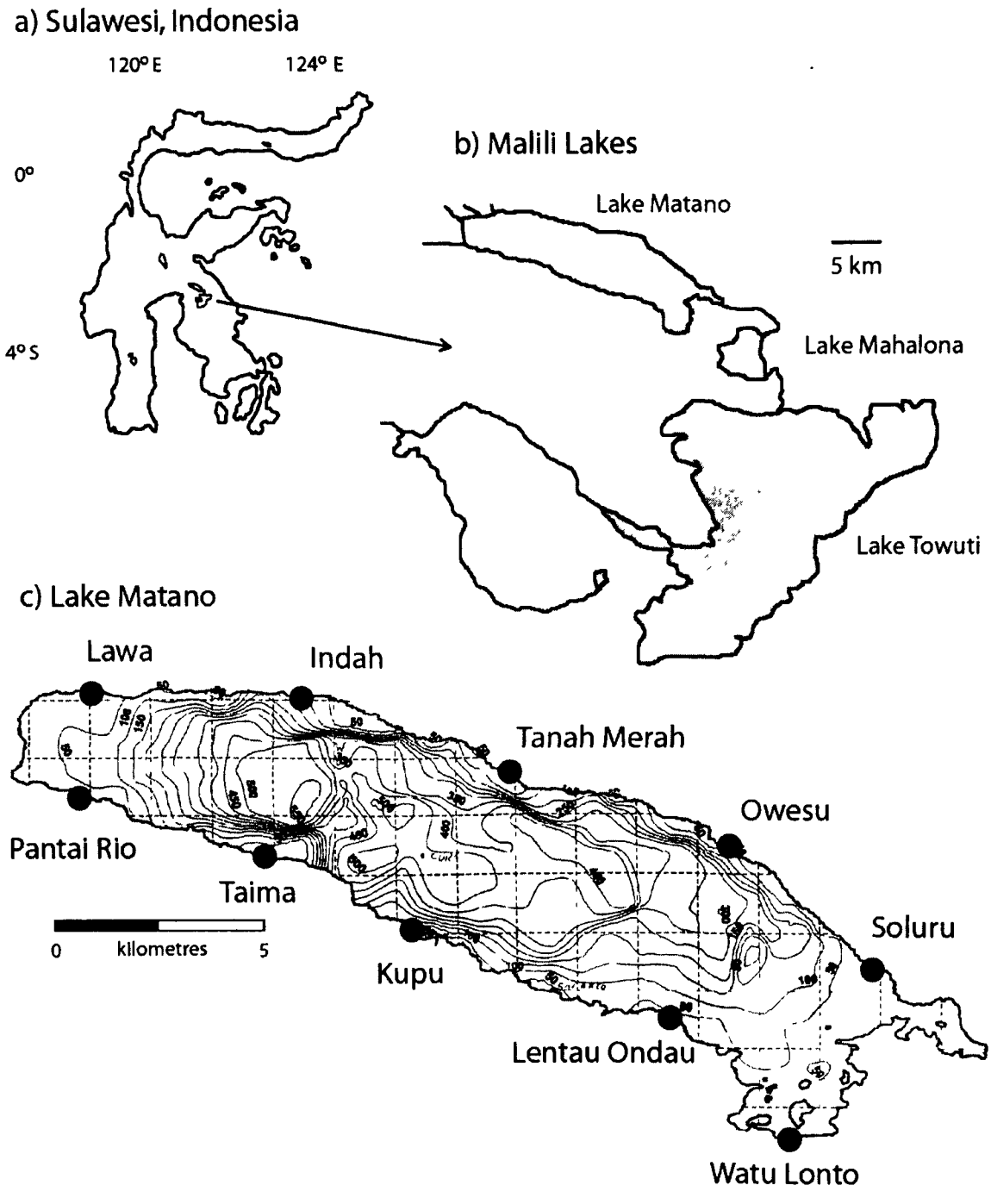


Figure 3.1 — Map of study area: a) Sulawesi, Indonesia. b) Malili Lakes in Southern Sulawesi. c) Lake Matano with sample sites indicated. Redrawn from Roy *et al,* (2007a).

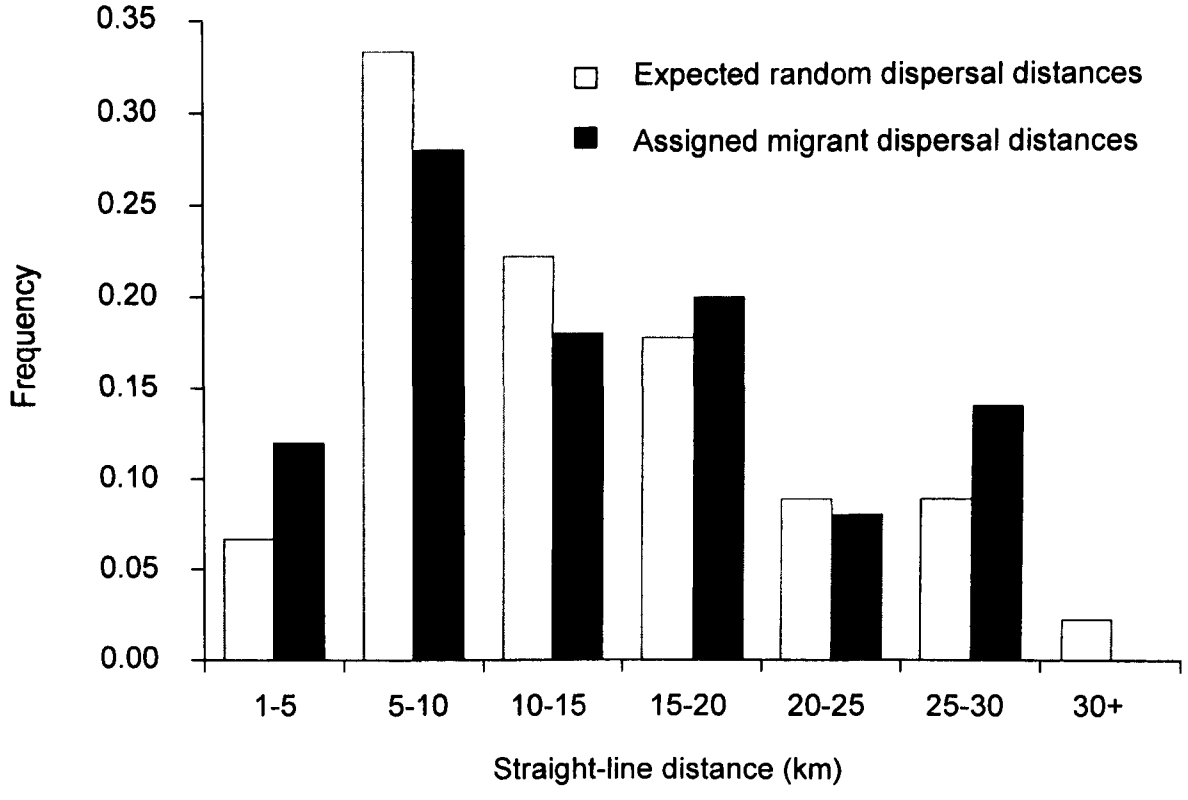


Figure 3.2 — Frequency histogram of open-water dispersal distances for individual migrant *Telmatherina antoniae* identified by genotype assignment with the comparable random expected distribution. The two distributions do not differ significantly (Kolmogorov-Smirnoff test ($P > 0.05$).

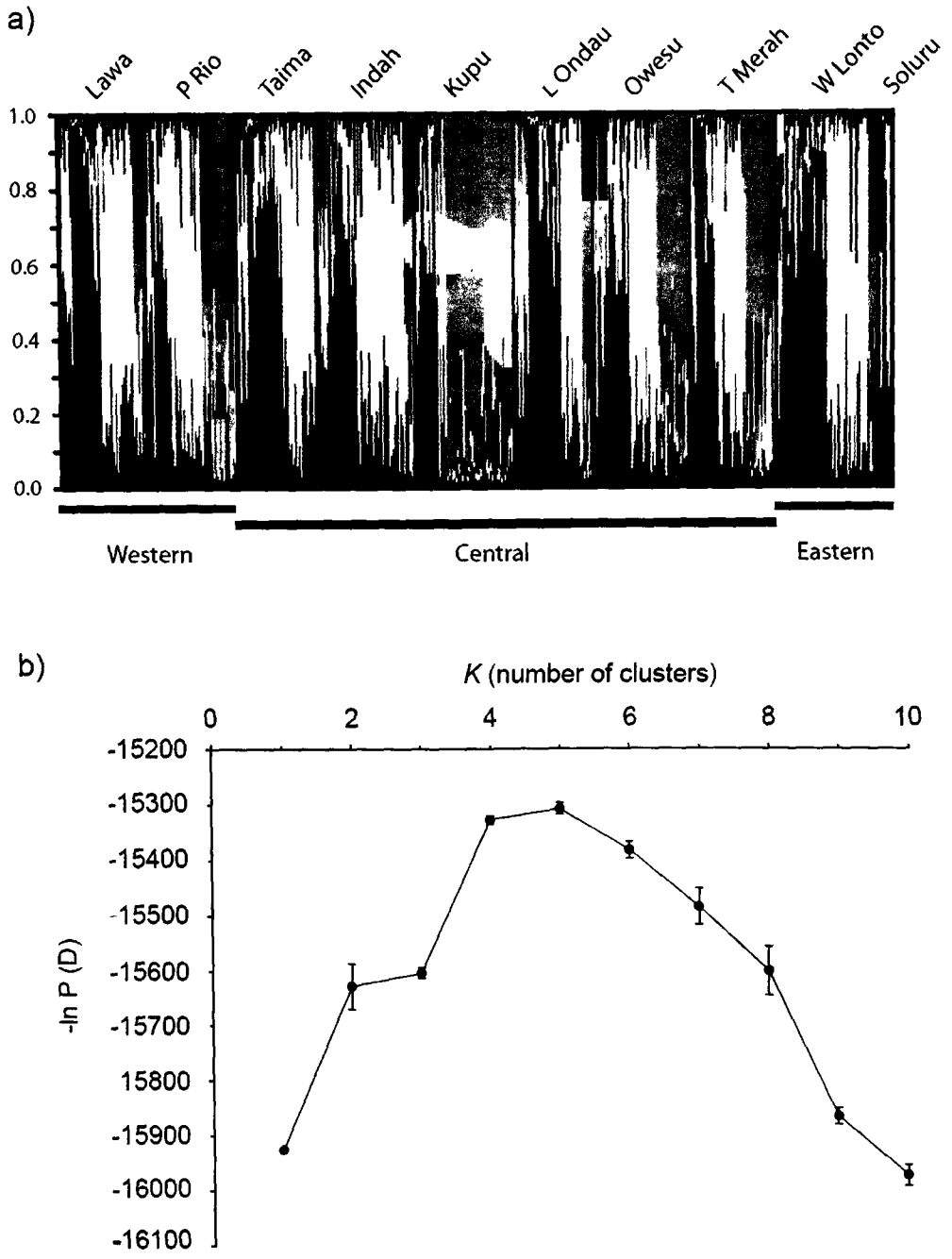


Figure 3.3 — a) Summary of membership proportions for individuals from geographic sites in Bayesian clusters using STRUCTURE ($K = 4$). The column shading corresponds to each of the four clusters. b) Negative log likelihoods (-ln P (D)) for microsatellite data fit to STRUCTURE models minimizing LD (k: 1-10), 500000 burn-in, 500000 iterations, with allele frequencies correlated and admixture allowed.

4.0 — NO BARRIERS TO GENE FLOW AMONG SYMPATRIC COLOUR-MORPHS OF "SMALL" *TELMATHERINA ANTONIAE* FROM INDONESIA'S LAKE MATANO*

4.1 INTRODUCTION

Since the formulation of the theory of natural selection, the maintenance of polymorphisms within a population has been a subject of great interest to biologists. The presence of colour polymorphisms within a species offers an opportunity for crypsis, predation avoidance, assortative mating among individuals and the potential for sexual selection for a particular colour morph. Sexual selection, through female mating preferences or male competitive behaviour, has been implicated as means of maintaining colour polymorphisms in natural populations (Kingston *et al.*, 2003; Gray & McKinnon, 2007). Since many species occupy variable habitats throughout their lifetimes, their relative conspicuousness to both potential mates and/or predators also varies (Gamble *et al.*, 2003; Chunco *et al*, 2007). Frequency- and density dependent selection for particular colour morphs have been shown to maintain colour polymorphisms in differing habitats and social environments (Sinervo *et al* , 2001; Svensson & Sinervo, 2004), even despite weak female preferences (Chunco *et al* , 2007). Thus, colour polymorphisms are rarely maintained by one mechanism and are likely sustained by a variety of ecological and behavioural factors including crypsis, predation, mate choice or competition for mates (Kingston *et al* , 2003; Gray & McKinnon, 2007).

* Walter RP, Haffner GD, Heath DD Genetic and behavioural data indicate no barriers to gene flow among sympatric colour-morphs of "small" *Telmatherina antoniae* from Indonesia's Lake Matano Second revision in review at *Journal of Fish Biology*

Among diverging but sympatric fishes, there is mixed evidence that colour polymorphisms play a role in the speciation process (Seehausen, 2000; Boughman, 2001; Danley & Kocher, 2001; Roy *et al.*, 2007). However, assortative mating, where individuals select mates of similar colour or patterning, has the potential to initiate genetic divergence via random drift and selection if gene flow is impeded between individuals of dissimilar colours. Such a process has been demonstrated in a number of fish species; notably the poeciliids (Endler, 1983; Kingston *et al.*, 2003) and African cichlids (Seehausen, 1997; Pauers *et al.*, 2004). Within the latter group there are striking examples of speciation via sexual selection, including the *in situ* isolation of Lake Victoria's sympatric *Pundamilia nyerei* Witte-Maas & Witte and *P. pundamilia* Seehausen & Bouton, two species with minimal morphological differentiation save colouration (Seehausen, 1997). Other species complexes show more regional variation in isolation among morphs. For example, in hamlets (*Hypoplectrus* sp., Serranidae), colour polymorphisms have been associated with assortative mating and sexual selection (Fischer, 1980; Graves & Rosenblatt, 1980; Michiels, 1998); however, reproductive isolation is not complete, as hybrid fertilization is still possible (Whiteman & Gage, 2007). Furthermore, localized ecological processes appear to influence genetic structure among certain colour morphs (Puebla *et al.*, 2008).

Within the Telmatherinid species complex of Lake Matano there are iridescent individuals of yellow, blue and mixed colour-morphs found in sympatry. Initial divergence in the species complex as a whole was shown to be more related to body-shape than primary body colour (Roy *et al.*, 2007b) with the group consisting of two major morphotypes based on fin-shape: "sharpfins" and "roundfins" (Herder *et al.*, 2006, Herder *et al.*, 2008) Within the "roundfins" is the species *Telmatherina antoniae* Kottelat which occurs in two genetically distinct

morphospecies—large and small (Herder *et al.* 2008)—both endemic to Lake Matano. A recent study of population genetic structure of *T. antoniae* revealed significant population genetic structure despite high dispersive capability, as a result of admixed proportions of at least four sympatric populations (Walter *et al.*2008). As *T. antoniae* are effectively an 'island' species, variation in visible genetic polymorphisms is expected to decrease due to genetic drift and inbreeding (Hayashi & Chiba, 2004).

In this study, genetic divergence among colour-morphs was assessed using polymorphic microsatellite loci to infer the occurrence of mating isolation along primary colour lines in the "small" *Telmatherina antoniae* morphospecies. As fierce inter-male fighting occurs in *T antoniae* (Gray & McKinnon, 2007), behavioural surveys were also used to measure the occurrence of assortative mating and to characterize inter-male aggression among paired males and intruders.

4.2 Materials and Methods

Given the status of *T. antoniae* as vulnerable IUCN Red-list threatened species (World Conservation Monitoring Centre, 1996; in IUCN, 2007), no individuals were sacrificed for this study A total of 428 individuals identified as "small" *T. antoniae* on the basis of morphology (following Kottelat, 1991; 1993) were sampled non-lethally via seine net from eight sites in Lake Matano, Indonesia (Fig 4.1). A small section of the anal fin tissue was removed and preserved in 95% ethanol and all fish were quickly released to their site of capture. DNA was recovered from tissue samples following the plate-based extraction method (Elphinstone *et al.*, 2003) and resuspended in 50 µl of Tris-EDTA buffer (10mM Tris, 1.0mM EDTA, pH 8.0). Confirmation of

species identification was made using mtDNA RFLP analysis following Roy (2006). On the basis of the colour of their body (males) and 2nd dorsal and anal fins (females), fish were sorted into three groups blue, yellow, or mixed. The samples were genotyped at ten microsatellite loci following Walter *et al.* (2007). The mean number of alleles, effective number of alleles, observed and expected heterozygosities and inbreeding coefficients (F) for each colour-morph from each site was calculated using GENALEX 6 (Peakall & Smouse, 2006). Fisher's Exact tests were performed to test for significant differences in allele frequency distributions among blue, yellow, and mixed colour morphs both within and among sites using GENEPOP (Raymond & Rousset, 1995). As the number of individuals of the 'mixed' individuals were generally low or absent altogether from some sampling sites, a test for significant genetic differentiation among colour morphs pooled from all sampling sites was performed using F_{ST} (Weir & Cockerham, 1984) in the program MSA 4.0 (Dieringer & Schlötterer, 2003). As significant spatial genetic differentiation has been found for this species (Walter *et al.*, 2008), hierarchical genetic differentiation was tested using AMOVA with colours nested within sampling sites using ARLEQUIN 3.0 (Excoffier *et al.*, 2005).

An independent heuristic approach for examining genetic structuring relying on Bayesian assignment was used to partition individuals into genetic clusters on the basis of an individual's multilocus genotype using BAPS 4.14 (Corander *et al.*, 2004). This analysis detects underlying genetic structure and determines whether this structure corresponds to the colour groups. Individual runs were performed for each sampling site; then data were pooled among all sites and the analysis rerun. BAPS clustering of individuals at individual sites was performed assuming 2, 3, 4 and 5 groups, with 3 replicates for each simulation; runs among pooled sites

were performed using the same scheme. Admixture analysis was then performed ignoring clusters with fewer than 5 individuals (n = 421, following removal of 7 individuals), and simulating 200 individuals in 50 iterations, to test for genetic mixture among these groups.

Inter-male aggressive behaviour was assessed by surveys of 33 courting male-female pairs, 17 of which the courting male was blue, and 16 of which the courting male was yellow, at Watu-Lonto sites along two 10m swimming transects within Lake Matano. Courting male-female pairs were observed for 3 min each, and when male-male aggressive displays were observed the following data was recorded: colour of courting male, colour of extra-pair male, and retention or loss of courted female. Contingency tables were used to test for assortative pairing along colour lines and for biases in the colour of extra-pair males. To maximize power in the analyses, male-female pairs were sorted into groups on the basis of 'like' and 'unlike' pairings (e.g., blue male-blue female pair = like).

4.3 RESULTS

Across pooled sites, 15 private alleles were found in the blue morphs, 11 in the yellow, and 3 among the mixed (Table 4.1). Fisher's exact tests and F_{ST} estimates revealed no significant genetic differentiation between blue and yellow morphs; with minor differentiation between yellow and mixed (Table 4.2). The hierarchical AMOVA with colours nested within sampling sites, revealed low but significant spatial differentiation, but no significant genetic variation among colour morphs (Table 4.3). The results of the Bayesian genotype clustering analysis indicated the presence of 3-5 genetic clusters per site, and 5 clusters among the pooled samples (Fig. 4.2.), however; none of these groups were associated with a particular colour-morph.

Courting *T. antoniae* included individuals of both primary colour morphs, but no mixed individuals. Surveys of 33 male-female pairs revealed 17 pairs of which males were blue and 16 pairs in which males were yellow (Table 4.4). Among the male-female pairs, six of the 17 blue-male pairs experienced two or more intruders during observation, while only three of 16 yellow male pairs experienced more than one intruder. The number of intruders a male experienced did not seem to have an effect on subsequent aggressive displays, and these data were included with all samples. The 2 x 2 contingency table showed that females were not randomly distributed with respect to the colour of the male, with 24 'like' pairings compared with 9 'unlike' pairings, $\chi^2_{0\,05,1} = 6.81$, P = 0.009. No significant colour biases were observed for aggressive displays toward extra-pair males, $\chi^2_{0\,05,\,1} = 1.88$, P = 0.170. Furthermore, the colour of the extra-pair male was not associated with loss or retention of a female.

4.4 DISCUSSION

A proposed hypothesis in studies of adaptive radiation of polychromatic species is that sexual selection based on female preferences for male colouration is a cause of rapid divergence (Lande, 1981; Panhuis *et al*, 2001; Knight & Turner, 2004). Despite evidence for genetic structure in *T. antoniae* at 10 microsatellite loci as indicated by both the AMOVA and Bayesian clustering analysis, no consistent patterns of genetic differentiation were drawn along primary colour lines. The only discrepancy was significant genetic differences observed among yellow and mixed morphs using both exact tests and F_{ST} estimates; however, a mechanism for such divergence among these morphs remains elusive. While mixed morphs in other Telmatherinids have been putatively suggested to be less 'fit' intermediate offspring of the primary colour

morphs, this cannot explain the observed genetic divergence. Rapid evolution of sperm recognition proteins has been implicated in the isolation of some fish species (Levitan & Ferrell, 2006); and could possibly account for the weak, but significant differentiation between yellow and mixed morphs.

The behavioural survey data did indicate significant assortative mating along primary colour lines. However, a small number of male-female pairs consisted of individuals with disimilar body colours, in agreement with Gray & McKinnon (2006). However, behavioural surveys showed that male aggression against intruding males was not biased against like colours, suggesting that males view conspecific males of different colour as an equal threat as those of similar colour. The perception of different colour forms as an equal threat is consistent with unpaired males exhibiting 'sneaky' mating tactics, regardless of their colour (Gray & McKinnon, 2006). Clearly, despite the significant positive assortative mating, sufficient cross-colour matings are occurring to result in a lack of genetic differentiation among the neutral loci employed here.

Examples of species exhibiting conspicuous colour polymorphisms with no underlying genetic differentiation at neutral loci are not rare; however such scenarios may occur as a result of individuals occupying different habitats throughout their lifetimes. For example, the remarkably clear waters of Lake Matano (secchi depth 20 m, Haffner *et al.*, 2001) likely facilitated the evolution of colour morphs via natural selection as particular colour morphs may be selected against in environments where they appear most conspicuous to predators (Ryan, 2001; Coyne & Orr, 2004). Gray *et al.* (2008) have suggested that differential reproductive success across visual environments is responsible for the occurrence of similar colour

polymorphisms in the closely-related *Telmatherina sarasinorum*, also from Lake Matano. According to this model, environment-contingent sexual selection occurs where males are favoured in habitats in which they are most conspicuous (Gray *et al.*, 2008). Additionally, an opportunity for sneaking among all Telmatherinid males does exist (Gray & McKinnon, 2007), which may supersede any assortative mating. Given the lack of genetic isolation at the microsatellite loci employed in this study, and the lack of genetic structure among colour morphs of *T antoniae* uncovered by Herder *et al.* (2008) using AFLP loci, its is likely that similar factors are at work in *T. antoniae* populations.

Studies of sexual selection tend to place a greater emphasis on female mate choice rather than choice by males or inter-male competition, when in fact the latter two may also contribute significantly. For example, inter-male competition frequently occurs in small aquaria and spatially-restricted artificial populations (Farr, 1977; Martin & Hengstebeck, 1981). For *T. antoniae*, inter-male competition appears to affect blue and yellow colour morphs more or less equally, but further assessment of inter-male competition for access to females across a variety of microhabitats, and their respective light regimes, is needed. Nonetheless, the lack of significant genetic differentiation between the primary colour morphs on a site-by-site basis likely reflects an equivalence of inter-male competition among colours.

This study shows that despite assortative mating on the basis of similar colour in Lake Matano *T. antoniae*, this assortative mating is not constraining gene flow among colour morphs. This surprising result leads to the question of how the colour polymorphisms are maintained in this endemic species. One possibility is environment-dependant sexual selection (known to occur

in a closely-related and sympatric species, Gray *et al.*, 2008) that may favour different colours in different parts of the lake. Alternatively, the colour variation may reflect habitat-dependant crypsis for predator avoidance. Finally, gene flow among the colour morphs may reflect recent reproductive barrier disruption, and the fish may be undergoing introgressive hybridization. Clearly additional study is warranted to determine how the *T. antoniae* colour morphs are maintained in Lake Matano.

4.5 REFERENCES

Allender, C. J., Seehausen, O., Knight, M.E., Turner, G. F., & Maclean, N. (2003). Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14074-14079.

Andersson, M. (1994). Sexual Selection. Princeton: Princeton University Press.

Boughman, J. W (2001). Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944-948.

Chunco, A.J., McKinnon, J.S., & Servedio, M.R. (2007). Microhabitat variation and sexual selection can maintain male color polymorphisms. *Evolution* **61**, 2504-2515. doi: 10.1111/j.1558-5646.2007.00213.x

Corander, J., Waldmann, P., & Sillanpaa, M.J. (2003). Bayesian analysis of genetic differentiation between populations. *Genetics* **163**, 367-374.

Coyne, J.A. & Orr, H.A. (2004). Speciation. Sinauer, Sutherland,.MA. 545 pp.

Danley, P. D. & Kocher, T.D. (2001). Speciation in rapidly diverging systems: lessons from Lake Malawi. *Molecular Ecology* **10**, 1075-1086. doi: 10.1046/j.1365-294X.2001.01283.x

Elphinstone, M.S., Hinten, G.N., Anderson, M.J. & Nock, C.J. (2003). An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Molecular Ecology Notes* **3**, 317-320. doi: 10.1046/j.1471-8286.2003.00397.x

Endler, J. A. (1983). Natural and sexual selection on colour patterns in poeciliid fishes. *Environmental Biology of Fishes* **9**, 173–190. doi:10.1007/BF00690861

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

Farr, J.A. (1977). Male rarity or novelty, female choice behaviour, and sexual selection in the guppy, *Poecilia reticulata* Peters (Pisces: Poeciliidae). *Evolution* **31**, 162–168.

Fischer, E.A. (1980). The relationship between mating system and simultaneous hermaphroditism in the coral reef fish, *Hypoplectrus nigricans* (Serranidae). *Animal Behavior* **28**, 620-33.

Gamble, S., Lindholm, A. K., Endler, J. A., & Brooks, R. (2003). Environmental variation and the maintenance of polymorphism: the effect of ambient light spectrum on mating behaviour and sexual selection in guppies. *Ecology Letters* **6**, 463–472. doi: 10.1046/j.1461-0248.2003.00449.x

Gorlick, D.L. (1976). Dominance hierarchies and factors influencing dominance in the guppy, *Poecilia reticulata* (Peters). *Animal Behaviour* **24**, 336–346.

Graves, J.E. & Rosenblatt R.H. (1980). Genetic Relationships of the Color Morphs of the Serranid Fish *Hypoplectrus unicolor. Evolution* **34**, 240-245.

Gray, S. M. & McKinnon, J. S. (2006). A comparative description of mating behaviour in the endemic Telmatherinid fishes of Sulawesi's Malili Lakes. *Environmental Biology of Fishes* **75**, 471-482.

Gray, S. M. & McKinnon, J. S. (2007). Linking color polymorphism maintenance and speciation. *Trends in Ecology & Evolution* **22**, 71-79.

Gray, S.M., Dill, L.M., Tantu, F Y., Loew, E.R., Herder, F & McKinnon, J.S. (2008). Environment contingent sexual selection in a colour polymorphic fish. *Proceedings of the Royal Society of London B* **275**, 1785-1791.

Haffner, G.D., Hehanussa, P.E. & Hartoto, D. (2001). The biology and physical properties of large lakes of Indonesia. In *The Great Lakes of the World: Food-Web Health and Integrity* (Manwar, M., Heck, R.E., eds), pp. 183-192. Leiden: Backuys.

Hayashi, M. & Chiba, S. (2004). Enhanced colour polymorphisms in island populations of the land snail *Euhadra peliomphala. Biological Journal of the Linnean Society* **81**, 417-425. doi: 10.1111/j.1095-8312.2003.00292.x

Herder, F., Pfaender, J. & Schliewen, U.K. (2008). Adaptive sympatric speciation of polychromatic "roundfin" sailfin silverside fish in Lake Matano. *Evolution* **62**, 2178-2195. doi: 0.1111/j.1558-5646.2008.00447.x

Kingston, J. J., Rosenthal, G. G. & Ryan, M. J. (2003). The role of sexual selection in maintaining a colour polymorphism in the pygmy swordtail, *Xiphophorus pygmaeus*. *Animal Behaviour* **65**, 735–743.

Knight, M.E. & Turner, G.F (2004). Laboratory mating trials indicate incipient speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra* from Lake Malawi. *Proceedings of the Royal Society of London B* **271**, 675–680.

Kottelat, M. (1991). Sailfin silversides Pisces: Telmatherinidae of Lake Matano, Sulawesi, Indonesia, with descriptions of six new species. *Ichthyological Exploration of Freshwaters* **1**, 321–344.

Kottelat, M., Whitten, A.J., Kartikasari, S.N. & Wirjoatmodjo, S. (1993). Freshwater Fishes of Western Indonesia and Sulawesi. Jakarta: Periplus.

Lande, R. (1981). Models of speciation by sexual selection on phylogenic traits. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 3721–3725.

Levitan, D.R. & Ferrell, D.L. (2006). Selection on gamete recognition proteins depends on sex, density, and genotype frequency. *Science* **312**, 267-269.

Martin, D. & Hengstebeck, M.F (1981). Eye color and aggression in juvenile guppies, *Poecilia reticulata*. *Animal Behaviour* **29**, 325–331.

Michiels, N.K. (1998). Mating conflicts and sperm competition in simultaneous hermaphrodites. In *Sperm Competition and Sexual Selection* (Birkhead, T R. &. Møller, A. P., eds), pp. 219–254, London: Academic Press.

Panhuis, T.M., Butlin, R., Zuk, M. & Tregenza, T. (2001). Sexual selection and speciation. *Trends Ecology & Evolution* **16,** 364–371.

Pauers, M.J., McKinnon, J.S., & Ehlinger, T.J. (2004). Directional sexual selection on chroma and within-pattern colour contrast in *Labeotropheus fuelleborni*. *Proceedings of the Royal Society of London B* **2004**, S444-S447.

Peakall, R. & Smouse, P.E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295. doi: 10.1111/j.1471-8286.2005.01155.x

Puebla, O., Bermingham, E., & Guichard, F (2008). Population genetic analyses of *Hypoplectrus* coral reef fishes provide evidence that local processes are operating during the early stages of marine adaptive radiations. *Molecular Ecology* **17**, 1405-1415. doi: 10.1111/j.1365-294X.2007.03654.x

Raymond, M. & Rousset, F (1995). Genepop (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

Roy, D. (2006). The evolutionary history and ecology of *Telmatherina* in Lake Matano: An example of adaptive radiation in an ancient lake. PhD Thesis, University of Windsor, Windsor, ON.

Roy, D., Paterson, G., Hamilton, P B., Heath, D. D., & Haffner, G. D. (2007). Resource-based adaptive divergence in the freshwater fish *Telmatherina* from Lake Matano, Indonesia. *Molecular Ecology* **16**, 35-48. doi: 10.1111/j.1365-294X.2006.03106.x

Roy, D., Docker, M.F., Haffner, G.D. & Heath, D.D. (2007b). Body shape vs. colour associated initial divergence in the *Telmatherina* radiation in Lake Matano, Sulawesi, Indonesia. *Journal of Evolutionary Biology* **20**, 1126–1137. doi: 10.1111/j.1420-9101.2006.01289.x

Ryan, M.J. (2001). Seeing red in speciation. *Nature* **411**, 900-901.

Seehausen, O. (1997). Distribution of and reproductive isolation among color morphs of a rock dwelling Lake Victoria cichlid. *Ecology of Freshwater Fish* **5**, 195-202. doi: 10.1111/j.1600-0633.1996.tb00133.x

Seehausen, O. (2000). Explosive speciation rates and unusual species richness in haplochromine cichlids effects of sexual selection. *Advances in Ecological Research* **31**, 237-274

Seehausen, O., van Alphen, J.J.M, & Witte, F (1997). Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808-1811.

Sinervo, B., Bleay, C. & Adamopoulou, C. (2001). Social causes of correlational selection and the resolution of a heritable throat colour polymorphism in a lizard. *Evolution* **55**, 2040–2052. doi:10.1111/j.0014-3820.2001.tb01320.x

Svensson, E. & Sinervo, B. (2004). Spatial scale and temporal component of selection in side-blotched lizards. *American Naturalist* **163**, 726–734.

Walter, R. P., Ovenden, J. R., & Heath, D. D. (2007). Characterization of four tetranucleotide and six dinucleotide microsatellite markers for use in *Telmatherina antoniae* and related species. *Molecular Ecology Notes* **7**. 651-653. doi:_10.1111/j.1471-8286.2006.01664.x

Walter, R.P., Haffner, G.D., & Heath D.D. (2008). Dispersal and population genetic structure of *Telmatherina antoniae*, an endemic freshwater Sailfin silverside from Sulawesi, Indonesia. *Journal of Evolutionary Biology*, doi:10.1111/j.1420-9101.2008.01645.x

Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Whiteman, E.A. & Gage, M.J.G. (2007). No barriers to fertilization between sympatric colour morphs in the marine species flock. *Journal of Zoology* **272**, 305-310.

Table 4.1 Summary of genetic data from 10 microsatellite loci. N = sample size, N_A = mean number of alleles per locus, A_P = number of private alleles, H_O = mean observed heterozygosity, H_E = mean expected heterozygosity, F = inbreeding coefficient.

Site	Colour	N	N_A	A_P	H_O	H_E	F
Lawa	Blue	18.8	9.2	5	0.56	0.72	0.24
	Mixed	3.0	3.8		0.60	0.64	0.15
	Yellow	27.8	10.7	2	0.58	0.73	0.21
Indah	Blue	17.0	8.4	2	0.62	0.69	0.11
	Mixed	1.0	1.6		0.60	0.30	-
	Yellow	37.0	11 1	2	0.57	0.71	0.22
Owesu	Blue	25.0	9.6	1	0.56	0.70	0.23
	Mixed	10.0	6.8	1	0.50	0.67	0.28
	Yellow	17.0	8.0	1	0.58	0.69	0.19
L Ondau	Blue	10.0	7.6		0.54	0.74	0.27
	Mixed	6.0	5.2	1	0.52	0.63	0.25
	Yellow	34.9	11.6	4	0.52	0.75	0.33
P Rio	Blue	32.0	10.2	1	0.54	0.67	0.24
	Yellow	10.0	6.6		0.62	0.65	0.09
Taima	Blue	16.0	8.8	2	0.58	0.71	0.22
	Mixed	3.0	3.4	1	0.57	0.59	0.04
	Yellow	28.0	10.8	1	0.56	0.74	0.28
T Merah	Blue	12.0	6.8		0.53	0.66	0.24
	Mixed	18.0	8.8	3	0.53	0.74	0.31
	Yellow	11.0	6.9		0.50	0.69	0.31
W Lonto	Blue	17.9	8.7	4	0.56	0.72	0.24
	Mixed	10.0	6.8		0.50	0.67	0.33
	Yellow	20.0	8.4	1	0.55	0.69	0.27

Table 4.2 F_{ST} estimates (below diagonal) and P-values from Fisher's Exact tests (above diagonal). Bold values indicate significant differentiation and * indicates a significant F_{ST} (P <0.05) following Bonferonni correction.

Colour	**Blue**	**Yellow**	**Mixed**
Blue	-	0.096	0.123
Yellow	0.003	-	**0.006**
Mixed	0.006	**0.014***	-

Table 4.3 Hierarchical AMOVA with colour-morphs nested within spatial sites.

Source of variation	d.f.	S.S.	Variance components	Percent of variance	*P*
Among sites	7	89.92	0.0751	2.14	0.000
Among colours within sites	14	61.67	0.0317	0.90	0.168
Within individuals	818	2782.07	3.4011	96.96	0.000
Total	839	2933.67	3.5078		
Fixation Indices					
F_{ST}	0.030				
F_{SC}	0.009				
F_{CT}	0.021				

Table 4.4 Summary of behavioural data for male-female pairs of *Telmatherina antoniae* colour-morphs along 10m swimming transects.

Male colour	*n*	# of fights with another male	
		Blue	Yellow
Blue male			
Like pair	13	7	8
Unlike pair	4	1	2
	17		
Yellow male			
Like pair	13	3	0
Unlike pair	3	3	1
	16		

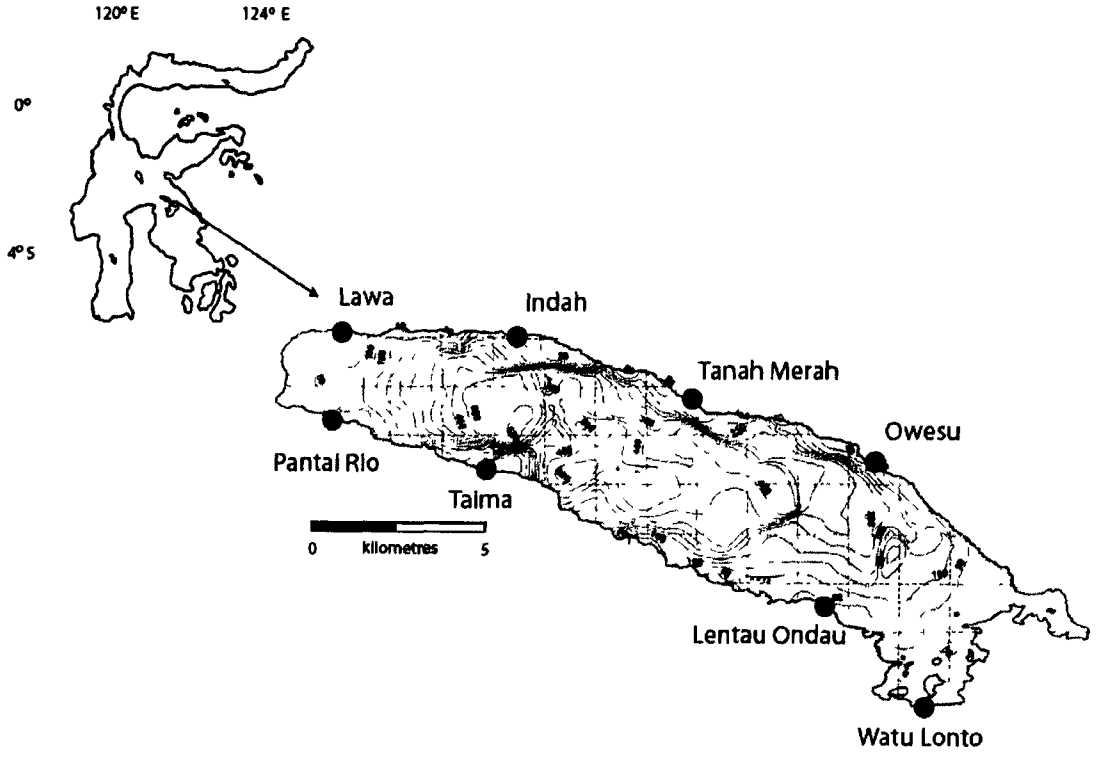


Figure 4.1 — Study and sampling sites in Lake Matano, Sulawesi, Indonesia.

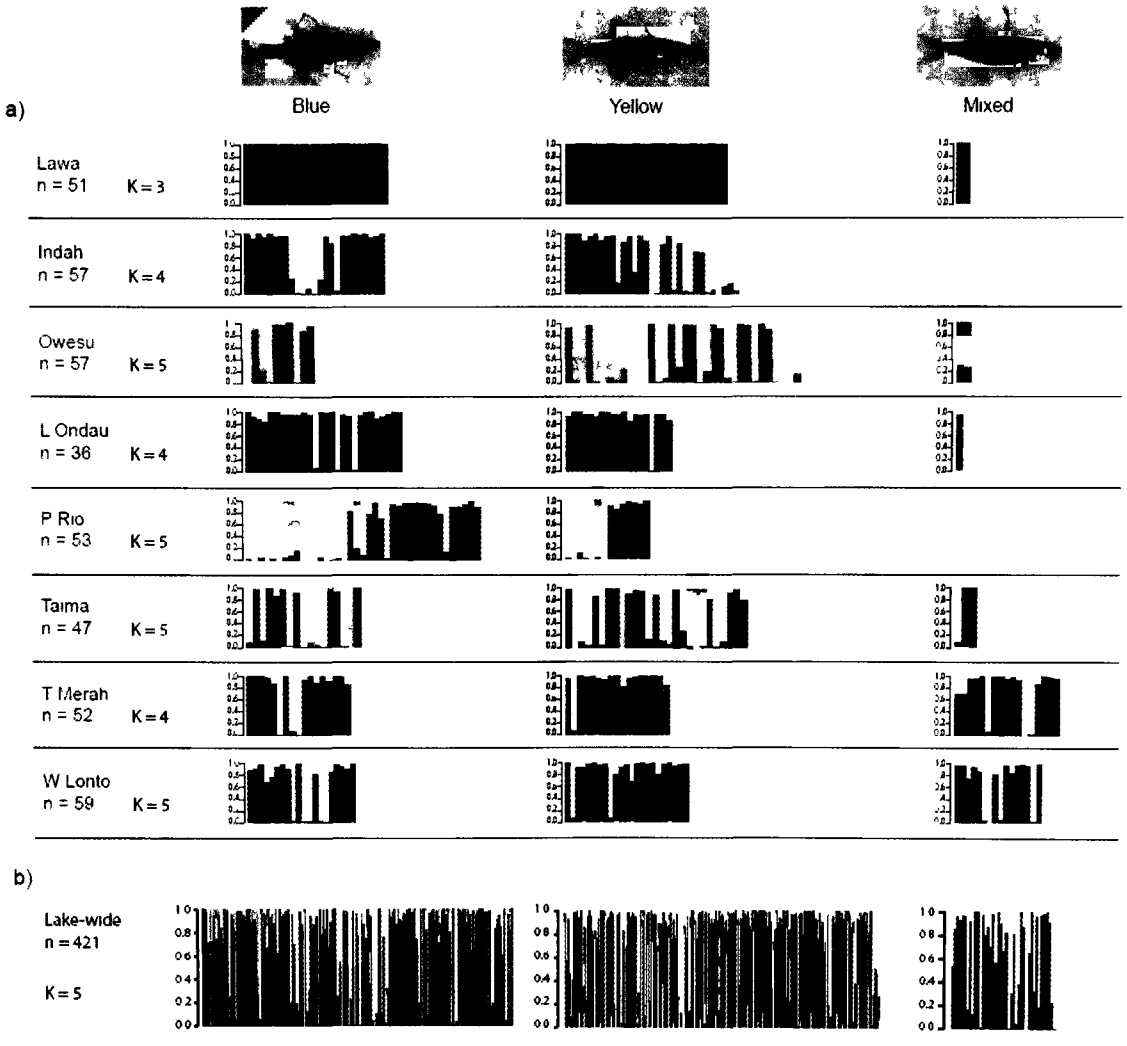


Figure 4.2 — Multilocus genotype clustering of *Telmatherina antoniae* colour-morphs using BAPS 4.14. Each bar corresponds to an individual fish in the blue, yellow or mixed groups; the colour of the bars represents the proportion of an individual fish's genetic assignment among most likely number of genetic clusters. a) Clustering performed by sampling site (K = 3-5). Cursory inspection may lead the reader to believe that variation in one given site exceeds that presented in the data from Chapter 3 (STRUCTURE); however, each site consists of an independent run. Therefore, colour codes on the bars representing an individual's genotype composition are not comparable across runs.; b) Clustering performed pooling all samples for a lake-wide estimate (K = 5). Photos by D. Roy, used with permission.

5.0 —ASYMMETRICAL RIVER DISPERSAL OF TWO ENDEMIC SAILFIN SILVERSIDES FROM SULAWESI'S SOUTHERN MALILI LAKES*

5.1 INTRODUCTION

Small populations are particularly sensitive to environmental perturbation as they often lack genetic diversity which buffers against extinction during periods of high mortality. Gene flow typically enhances genetic diversity; however geographic barriers to dispersal ultimately shape the biogeographic range of a species (Bohonak 1999). In aquatic systems, physical barriers such sand bars, deep water, river outflows and discontinuous habitat influence spatial genetic structure, as populations reflect environmental heterogeneity (Danley et al. 2000; Rico &Turner 2002). Rivers, however, can simultaneously act as barriers to dispersal and as effective transport vectors. For example, rivers often facilitate the movement of individuals primarily in the direction of flow, yet features such as distance, meanders, rapids, and sharp declines in elevation may serve to either limit or prevent upstream travel. In this manner, rivers facilitate asymmetrical gene flow from upstream to downstream populations.

The patterns of endemism in the fishes of the Malili Lakes region of Sulawesi, Indonesia, suggest that rivers act as both barriers and pathways to dispersal and gene flow (Roy 2006; Herder et al. 2006a; Schwarzer et al. 2008). The Malili Lakes (Fig. 5.1) constitute a series of three major lakes partially isolated from one another (Kottelat 1990). Lake Matano forms the headwaters of the Malili Lakes (Brooks 1950) and hosts a number of endemic fish species (Kottelat 1990; Roy 2006). The fish species of the lower Malili Lakes have distributions that

* Walter RP, Hogan JD, Haffner GD, Heath DD Asymmetrical river dispersal of two endemic Sailfin silversides from Sulawesi's southern Malili Lakes *Manuscript submitted to Conservation Genetics Jan 2009*

cover a larger area than those of Lake Matano. This distribution appears to be the result of successful downstream dispersal along the Tominanga River joining Lake Mahalona with Lake Towuti. This distribution of *Paratherina* is interesting as the Petea River, joining the headwaters of Lake Matano with Lake Mahalona has been previously described as a barrier to downstream dispersal (Roy 2006), leaving *Paratherina* endemic to the lower Malili Lakes.

Lake Mahalona is less than a tenth the size of Lake Towuti; however, both lakes have active local fisheries for species of *Paratherina* and the related Telmatherinid genus *Tominanga*. Given the active fisheries by the local community, an understanding of the underlying connectivity among lake populations has important implications for the successful conservation and management of these species. Effective species management relies on knowledge species philoptary, expected among-population movement patterns, and the consequences of movement for genetic structure (Crandall et al. 2000, Fraser and Bernatchez 2001, Schlosser et al. 2008).

This paper examines dispersal and genetic connectivity between two lake populations of *Paratherina* species using multilocus genetic methods. First, using polymorphic microsatellite markers I show contemporary genetic connectivity between Lakes Mahalona and Towuti. Secondly, I test whether the observed gene flow between lake populations is asymmetrical as a result of predominantly down-river dispersal. I observed similar patterns in two related sister species of *Paratherina*, and provide support for the Tominanga River as an asymmetrical vector for genetic connectivity within the southern Malili Lakes. The conservation of endemic tropical lake fish species faces severe logistical challenges, yet quantification of patterns of inter-lake

connectivity provides critical information for setting priorities and establishing guidelines for management action.

5.2 METHODS

Study species

Sailfin silversides (Telmatherinidae) are a family of tropical fishes endemic to the freshwaters of Sulawesi and brackish and mangrove swamps of New Guinea (Saeed and Ivanstoff, 1991; Kottelat, 1993). *Paratherina* species are restricted to the waters of Lakes Mahalona and Towuti (Kottelat 1993, Herder 2006). Juveniles and adults form small schools (4-10 individuals) at depths ranging from 1-3 m (Kottelat 1990). *Paratherina* species were originally described from Lake Mahalona by Aurich (1935), but have been observed in both Lakes Mahalona and Towuti (Kottelat 1990; Herder et al. 2006).

Study sites and sample collection

Approximately 100 fish per site were collected using beach seine and light trap netting from three sites in Lake Mahalona and three locations in Lake Towuti in May-June 2005 and June 2006. At the time of collections, specimens were identified as *Paratherina* spp 1 (*wolterecki*) and 2 (*striata*), however species designations were tentative. All specimens were photographed and fin-clipped before being released back into the lake. Collected fin tissue was stored in 95% ethanol.

DNA extraction and microsatellite genotyping

DNA was extracted from fin clip samples following the plate-based extraction method (Elphinstone *et al.* 2003). All DNA was resuspended in 50 µl of Tris-EDTA buffer (10mM Tris, 1.0mM EDTA, pH 8.0). Fish were genotyped at seven microsatellite markers developed for use in *Telmatherina antoniae* (Tan12, Tan13, Tan14, Tan16, Tan17.1, Tan24, Tan26; Walter et al. 2007). PCR amplifications were: initial denaturation at 94° C for 2 min; then 24 cycles at 94° C for 10 sec, annealing at primer specific temperature for 10 sec and extension at 72° C for 30 sec; and a final extension at 72° C for 1.5 min (Walter et al. 2007). The size of the PCR products was estimated using a LiCor 4300 DNA Analyzer with GENEIMAGIR 4.05 software (Scanalytics, Inc).We used MSAv4.0 (Dieringer & Schlötterer 2003) to test for genotyping errors, MICROCHECKERv2.2.3 (van Oosterhout 2004) to check for the presence of null alleles, and GENEPOP4.0 (Raymond & Rousset 1995) to test for Hardy-Weinberg (HWE) and linkage disequilibrium among loci.

Genetic analyses

Observed and expected heterozygosities and the inbreeding coefficient F were estimated in GENALEX v6.1. We then calculated F_{ST} estimates (θ, Weir & Cockerham (1982)) in GENEPOP to determine if significant genetic differentiation was observed between sites within and among lakes.

The direction of gene flow was investigated using a coalescent-based maximum-likelihood strategy in the program MIGRATE (Beerli 1998; Beerli and Felsenstein 1999). We performed the MIGRATE analyses in three runs to allow convergence of parameters. All runs were performed using the Brownian motion microsatellite model, using 10 000 burn-in runs and

the 'adaptive' heating setting, where four Markov chains are run at different 'temperatures': 1, 1.3, 3, 5. The parameters Θ and M represent N_e and migration rate, respectively, and the initial run allowed the estimation of Θ and M from F_{ST} estimates. The 2nd and 3rd run used Θ and M values obtained from the previous run entered as priors. Convergence of both Θ and M was observed in the 3rd run.

To further assess connectivity and gene flow direction we used genotype assignment in GENECLASS 2.0 (Piry et al. 2001). Individuals were assigned to either Lake Mahalona or Lake Towuti using the frequencies-based methods of Cornuet et al. (1995) and the 'leave-one-out' procedure. As genetic differentiation was strongest among lakes rather than sites within a lake, we pooled samples by lake and ran assignments with the goal of assigning individuals to a lake based on their multilocus genotype. A small number of fish (N = 4) caught in the Tominanga River < 1 km downstream from M2 were identified as *Paratherina striata* and assigned back to M2. Additionally, three individuals tentatively identified as *Paratherina wolterecki* sampled 1 km upstream from T1 in the Tominanga River, were assigned back to T1.

5.3 RESULTS

All sampling sites conformed to Hardy-Weinberg equilibrium across all loci following Bonferroni correction. High allelic diversity was present at all loci, with the average number of alleles per locus ranging from 7.1 to 19.3 (Table 5.1). For *Paratherina striata*, the mean number of alleles per locus also differed between Lake Mahalona and Lake Towuti populations, but *F*-values were comparable across all populations (Table 5.1). All pairwise F_{ST} estimates were significantly different from zero (Table 5.2). For *Paratherina wolterecki*, inbreeding coefficients

(F) differed between Lake Mahalona and Lake Towuti populations; the average number of alleles per locus dropped by five.

Estimates of directionality of gene flow revealed asymmetrical gene flow for all Mahalona-Towuti site pairwise comparisons (Fig. 5.2). Downstream dispersal events were found to be 2-3 times more common than upstream dispersal events in both species.

Self-assignment rates were close to 80% in both species. The relative power (D_{LR}) of genotype assignment is displayed as a plot of –ln log likelihoods (Fig. 5.3). Genotype assignment revealed slight asymmetrical dispersal patterns as higher numbers of migrant fish in Lake Towuti than in Lake Mahalona for both species. For *Paratherina wolterecki*, 16.3% of individuals sampled in Mahalona were assigned Towuti, compared with 19.2% of the Towuti samples were assigned back to Mahalona. Similarly, for *Paratherina striata* 17% of Mahalona individuals were assigned to Towuti, and 27% of the fish caught in Towuti were assigned back to Mahalona.

5.4 DISCUSSION

This study demonstrates the importance of rivers as vectors of connectivity among lake-dwelling populations. These results reveal contemporary genetic connectivity among fish populations endemic to the lower Malili lakes, based on results for two species of Sailfin silversides. The asymmetrical pattern of riverine gene flow in both species supports the large role of river current direction in driving dispersal in these small tropical fish. Based on random dispersal and expected population size, we would have predicted higher gene flow from the larger Lake

Towuti to Lake Mahalona (10 times larger based on surface area); however, the Tominanga River clearly acts to bias dispersal downstream. This bias in gene flow appears to be related to the flow regime of the Tominanga River but could also reflect the swimming capabilities of *Paratherina* species, as fish would be expected to more easily pass downstream.

Endemic species with restricted ranges pose serious conservation challenges due to global anthropogenic disturbance impacts on biodiversity (Gaston, 1996; Channell and Lomolino, 2000). Neighbouring populations can act as genetic diversity sources under metapopulation dynamics, thus isolated endemic species face elevated risk for inbreeding, loss of genetic variability, and ultimately, risk of extinction. For the *Paratherina* species of Sulawesi, we identified that populations in Lake Mahalona tend act as net exporters of migrants into Towuti. Furthermore, Lake Towuti populations, given the size of Lake Towuti, are not likely to be highly dependent on gene flow from Lake Mahalona. Moreover, Lake Mahalona is not likely to be 'rescued' by migrants from Lake Towuti. Thus, this information is both critical and timely in light of any environmental and anthropogenic perturbation affecting geographically-restricted organisms.

Isolated endemic species are characterized by small (effective) population size and are thus likely to lose rare alleles and display reduced genetic diversity through genetic drift and inbreeding (Daniels et al. 2000; Sherwin and Moritz 2000), which can, in turn, compromise fitness and population viability (Vrijenhoek 1994; Keller et al. 1994; Saccheri et al. 1998; Keller and Waller 2002; Palstra and Ruzzante 2008). Effective conservation relies on the identification of 'evolutionary significant units' or primary population subdivisions (Frasier and Bernatchez

2001; Olivieri et al. 2008). Population genetic structure for both *Paratherina* species is obvious given the marked genetic differentiation (F_{ST}, Table 2) and D_{LR} (Fig. 3). While there is effective gene flow between these two lakes for both species, it is not pronounced enough to erode neutral genetic differentiation occurring in either lake. In most populations, one effective migrant per generation is in principle enough to avoid the effects of genetic drift, however when drift is exacerbated by a low effective population size to census population size ratio, more migrants may be needed to maintain genetic diversity (Mills and Allendorf 1996; Wang 2004).

Understanding the relationship between dispersal and gene flow is central to effective wildlife management (Schlosser et al. 2008). This study highlights the importance and complexity of rivers as vectors of connectivity between lacustrine populations. Anthropogenic disturbances including mining and road construction within the watershed and fishing pressures within the lakes themselves highlights the need for baseline data for endemics in this biodiversity hotspot. Populations of *Paratherina* in Lake Mahalona should probably be considered as separate evolutionary significant units (ESAs) on the basis of geographic isolation, restricted gene flow and the potential of locally adapted phenotypic traits (Conner & Hartl 2004). The Lake Mahalona basin is geographically isolated from L Towuti, connected only by the Tominanga River. This river appears to be the only physical connection to other lacustrine *Paratherina* populations within the Malili System. The significant genetic differentiation within lake populations highlights the potential for local adaptation, while the between lake populations of *Paratherina* suggests that Lake Mahalona exports more migrants than it receives. Although ultimately dependent on effective population size, populations in Lake Mahalona may be at higher risk of extinction if exposed to significant environment perturbation.

5.5 REFERENCES

Beneteau C, Mandrak N, Heath DD (2008) The effects of river barriers and range expansion of the population genetic structure and stability in Greenside Darter (*Etheostoma blennioides*) populations. Conserv Genet doi: 10.1007/s10592-008-9627-9

Beerli P (1997–2004) Migrate: Documentation and Program, Part of lamarc, Version 2.0. Revised December 23, 2004. Available from URL: http://evolution.gs.washington.edu/lamarc.html.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics, 152, 763–773.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. Q Rev Biol 74: 2145.

Brooks JL (1950) Speciation in ancient lakes concluded. Q Rev Biol 25: 131-176.

Channel R, Lomolino M (2000) Dynamic biogeography and conservation of endangered species. Nature 403: 84–86.

Conner J, Hartl D (2004) *A Primer of Ecological Genetics*. Sinauer, Sunderland, MA

Cornuet J-M, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude population as origins of individuals. Genetics, 153, 1989–2000.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends Ecol Evol 15:290–295. doi:10.1016/S0169-5347(00) 01876-0

Daniels SJ, Triddy JA, Walters JR (2000) Inbreeding in small populations of red-cockaded woodpeckers: Insights from a spatially explicit individual-based model. In: Genetics, Demography and Viability of Fragmented Populations (eds. Young AG, Clarke GM), pp. 129–148. Cambridge University Press, Cambridge.

Danley PD, Markert JA, Arnegard ME, Kocher TD (2000) Divergence with gene flow in rock-dwelling cichlids of Lake Malawi. *Evolution* 54: 1725-1737

Elphinstone MS, Hinten GN, Anderson MJ, Nock CJ (2003) An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. Mol Ecol Notes 3:317–320. doi:10.1046/j.1471-8286.2003.00397.x

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Mol Ecol 10:2741–2752

Fraser DJ, Lippe C, Bernatchez L (2004) Consequences of unequal population size, asymmetric gene flow and sex-biased dispersal on population structure in brook charr (Salvelinus fontinalis). Mol Ecol 13: 67–80.

Fraser DJ, Hansen MM, Østergaard S, Tessier N, Legault M, Bernatchez L (2007) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. Mol Ecol 16:3866–3889. doi:10.1111/j.1365-294X. 2007.03453.x

Gaston, KJ (1996) Species-range-size distributions: patterns, mechanisms and implications. TREE 11: 197–201.

Herder F, Schwarzer J, Pfaender J, Hadiaty RK, Schliewen UK (2006) Preliminary checklist of sailfin silversides (Pisces: Telmatherinidae) in theMalili Lakes of Sulawesi (Indonesia), with a synopsis of systematics and threats. Verh Ges Ichthyol 5: 139–163.

Keller LF, Arcese P, Smith JNM, Hochachka WM, Stearns SC (1994) Selection against inbred song sparrows during a natural population bottleneck. Nature, 372, 356–357

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. TREE, 17, 230–241.

Mills LS, Allendorf FW (1996) The one-migrant-per-generation rule in conservation and management. Conserv. Biol., 10, 1509–1518.

Palstra FP, Ruzzante DE (2008) Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Mol Ecol 17: 3428-3447

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6, 288-295. doi: 10.1111/j.1471-8286.2005.01155.x

Piry S, Alapetite A, Cornuet JM et al (2004) GENECLASS 2: a software for genetic assignment and first generation migrants detection. J Hered 95: 536–539. doi:10.1093/jhered/esh074

Oliveiri LR, Hoffman JI, Higst-Zaher E, Majhuf P, Muelbert MMC, Morgante S, Amos W (2008) Morphological and genetic evidence for two evolutionarily significant units (ESUs) in the South American fur seal, *Arctocephalus australis*. Conserv Gen. 9:1451–1466. doi: 10.1007/s10592-007-9473-1

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA 94: 9197–9201. doi:10.1073/pnas.94.17.9197

Raymond M, Rousset F (1995) An exact test for populations differentiation. Evol Int J Org Evol 49: 1280–1283. doi:10.2307/2410454

Rico C, Turner GF (2002) Extreme microallopatric divergence in a cichlid species from Lake Malawi. Mol Ecol 11: 243-251.

Roy D (2006) The evolutionary history and ecology of *Telmatherina* in Lake Matano: An example of adaptive radiation in an ancient lake. PhD Thesis, University of Windsor, Windsor, ON.

Roy D, Kelly DW, Fransen CHJM, Heath DD, Haffner GD (2006) Evidence of small scale vicariance in *Caridina lanceolata* (Decapoda: Atyidae) in the Malili Lakes, Sulawesi Island, Indonesia. Evol Ecol Res 8: 1087-1099.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. Nature, 392, 491–494.

Schlosser JA, Dubach JM, Æ Trenton W. J. Garner TWJ, Araya B, Bernal M, Simeone A, Smith KA, Wallace RS (2008) Evidence for gene flow differs from observed dispersal patterns in the Humboldt penguin, Spheniscus humboldti Conserv. Genet. DOI 10.1007/s10592-008-9644-8

Sherwin WB, Moritz C (2000) Managing and monitoring genetic erosion. In: Genetics, Demography and Viability of Fragmented Populations (eds. Young AG, Clarke GM), pp. 9–34. Cambridge University Press, Cambridge.

Schwarzer J, Herder F, Misof B, Hadiaty RK, Schliewen UK (2008) Gene flow at the margin of Lake Matano's adaptive sailfin silverside radiation: Telmatherinidae of River Petea in Sulawesi. Hydrobiologia 615: 201-213.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4: 535–538. doi:10.1111/j.1471-8286.2004.00684.x

Vrijenhoek RC (1994) Genetic diversity and fitness in small populations In: Conservation Genetics (eds. Loeschcke V, Tomiuk J, Jain SK), pp. 37–53. Birkhäuser Verlag, Basel.

Wang J (2004) Application of the one-migrant-per-generation rule to conservation and management. Conserv. Biol., 18, 332–343.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evol 38: 1358–1370. doi:10.2307/2408641

Table 5.1 Summary of microsatellite genetic data for two *Paratherina* species from Lakes Mahalon and Towuti, Indonesia. N = average sample size, N_a = number of alleles, H_o = observed heterozygosity, H_e = expected heterozygosity, F = inbreeding coefficient.

Site	N	N_a	*Ho*	*He*	*F*
Paratherina wolterecki					
M1	85.3	15.4	0.793	0.847	0.064
Pet	50.9	14.9	0.785	0.828	0.055
T1	28.7	10.0	0.760	0.802	0.052
T2	25.9	12.4	0.846	0.834	-0.011
Paratherina striata					
M2	80.0	15.9	0.730	0.814	0.110
T1-2	51.3	15.4	0.799	0.824	0.026
T2-2	67.9	19.3	0.823	0.874	0.063
T3-2	59.9	16.1	0.723	0.840	0.151

Table 5.2 Pairwise F_{ST} estimates (Weir and Cockerham, 1983) for two *Paratherina* species sampled from Lakes Mahalona and Towuti, Indonesia. M1 = Mahalona 1 ; M2 = Mahalona 2; Pet = Petea outflow (Mahalona); T1 = Towuti 1; T2 = Towuti 2; T3 = Towuti 3. All values significant (P<0.05, except bold values).

Paratherina wolterecki				
	M1	Pet	T1	T2
Pet	0.024			
T1	0.048	0.051		
T2	0.026	0.027	0.025	

Paratherina striata				
	M2	T1	T2	T3
T1	0.032			
T2	0.021	0.018		
T3	0.013	**0.010**	**0.006**	

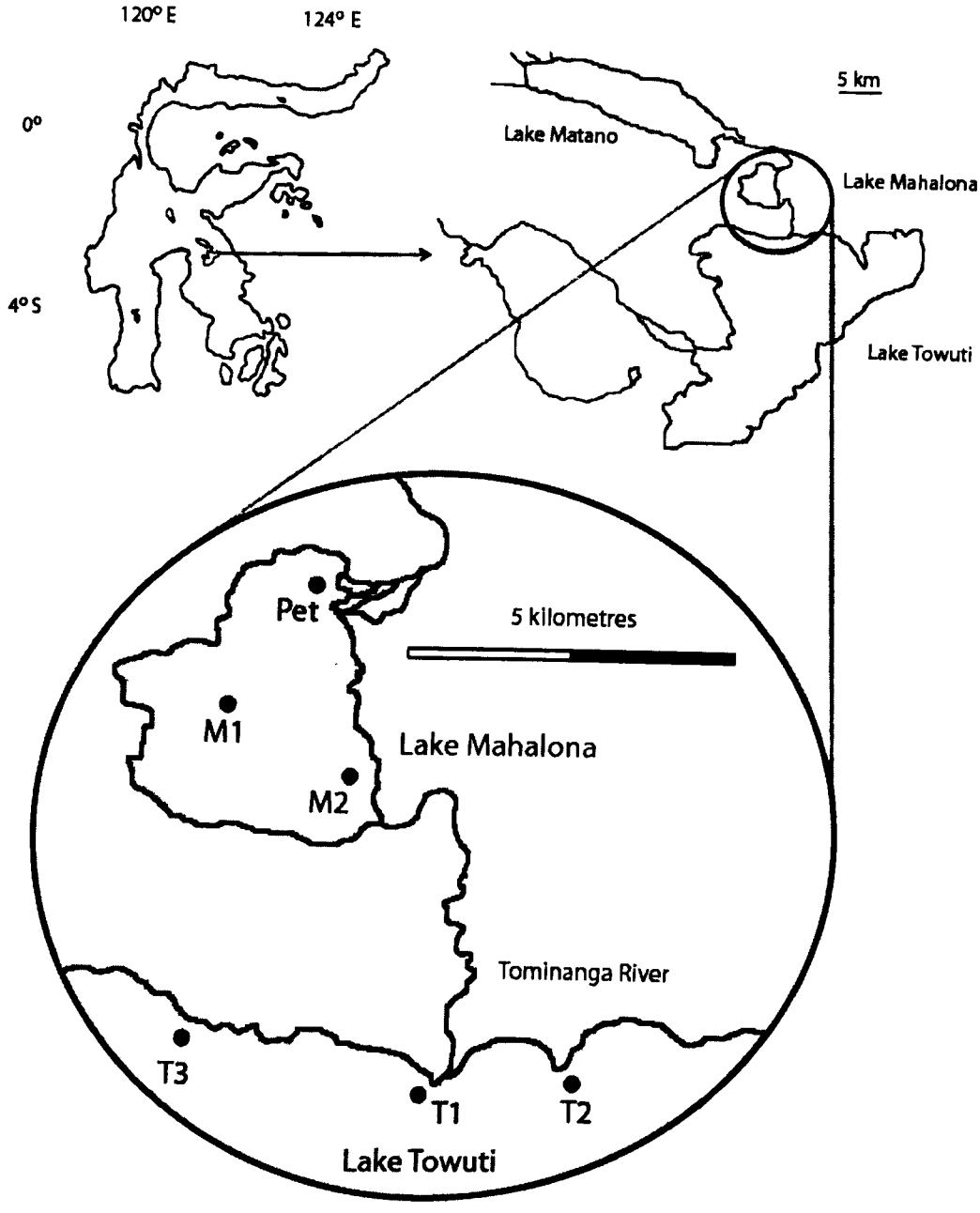


Figure 5.1 — Map of Malili Lakes, Sulawesi, Indonesia with sampling sites for *Paratherina wolterecki* and *P striata*.M_{1-2}= Mahalona, Pet = Petea River outflow, T_{1-3} = Towuti sites.

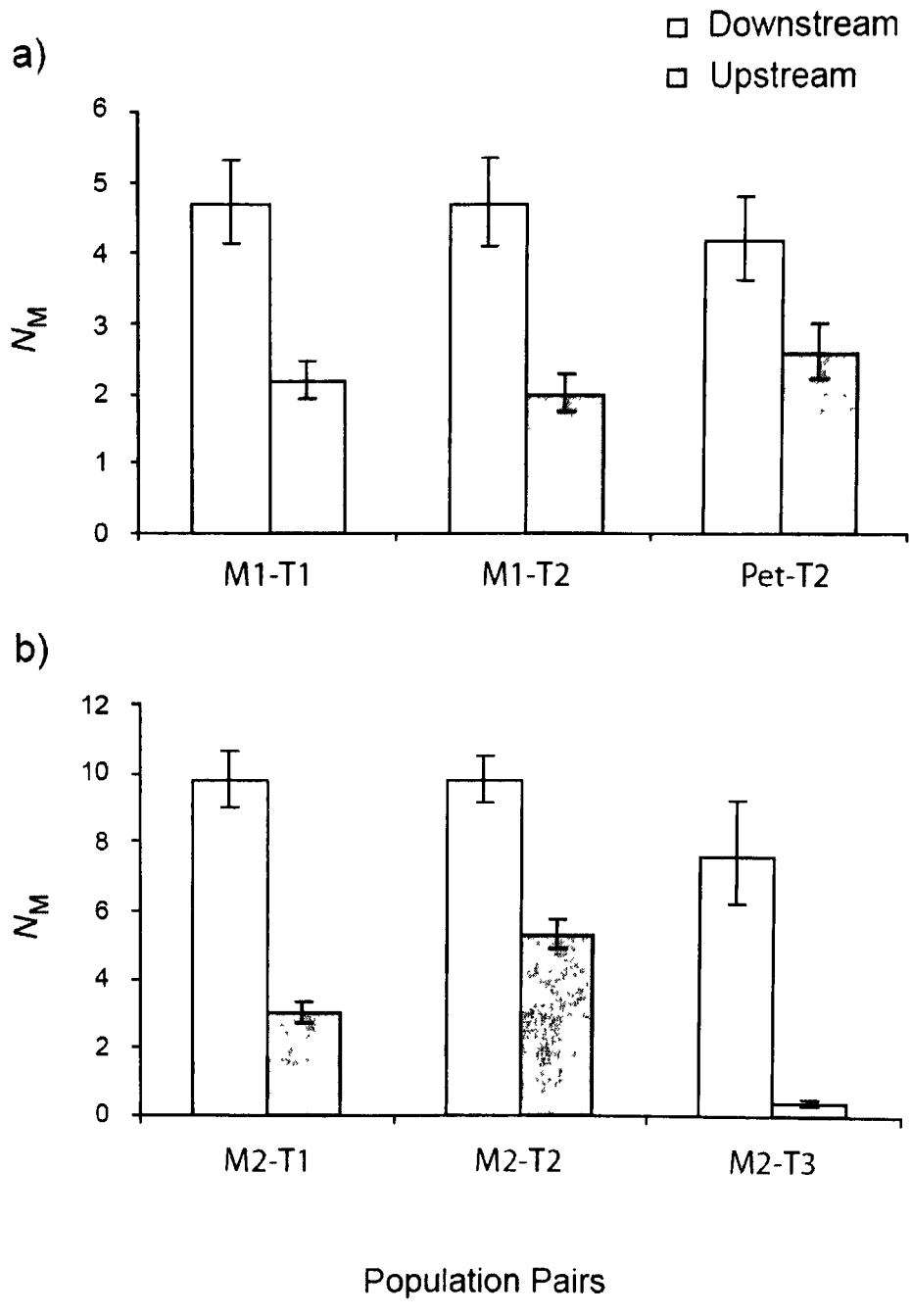


Figure 5.2 — Estimated number of migrants (N_M) for a) *Paratherina wolterecki* and b) *Paratherina striata* moving upstream versus downstream along the Tominanga River connecting the two sampled lakes (Lakes Mahalona and Towuti) using coalescent-based maximum-likelihood approach of Beerli & Felsenstein (1999). N_M estimates are for population pairs. See Figure 1 for locations of populations.

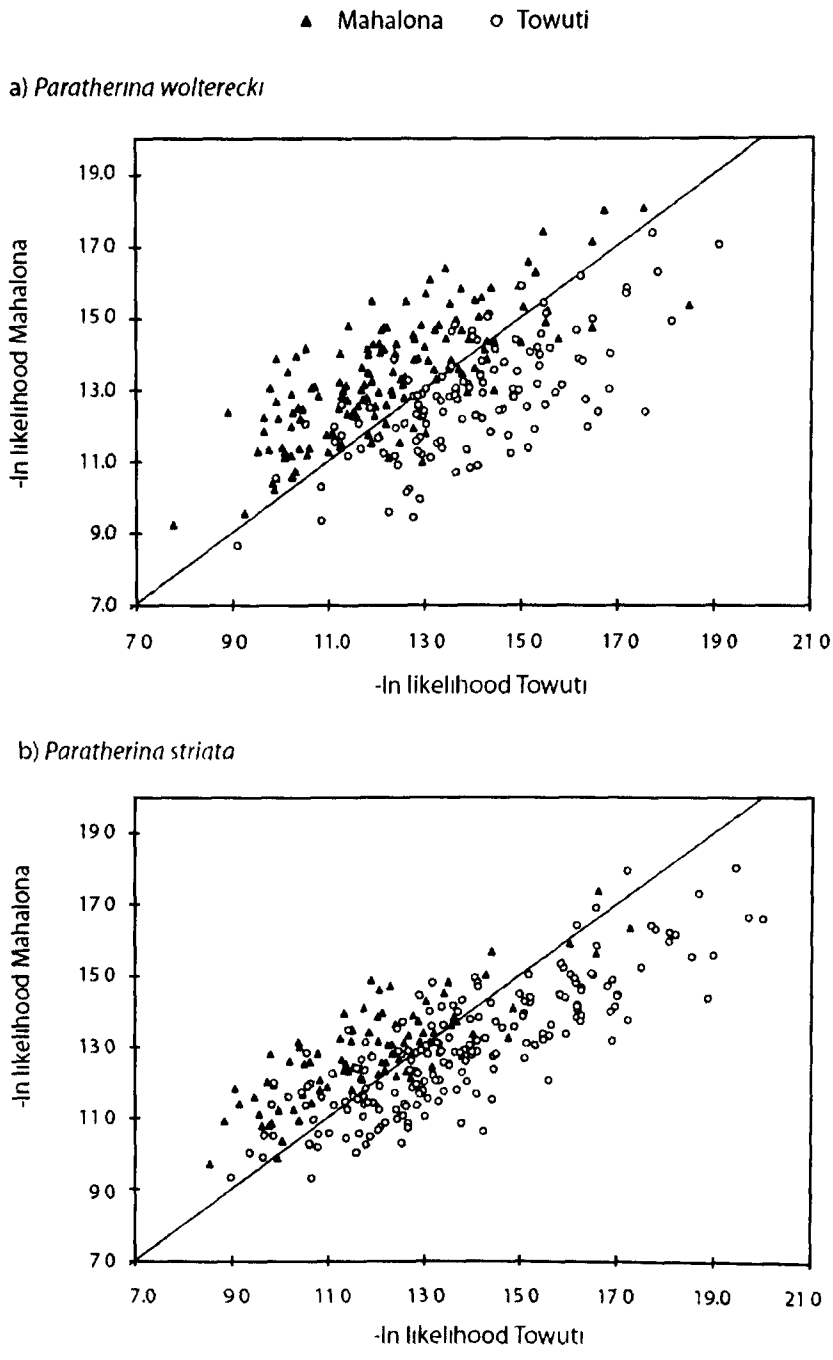


Figure 5.3 — Relative 'discriminatory power' of the assignment test using the criterion of Cornuet et al. (1995). The further the 'clouds of points of a given lake are from one another, the greater the power of the assignment method.

6.0 — PARALLEL ECOLOGICAL DIVERGENCE IN SULAWESI'S SAILFIN SILVERSIDE RADIATION

6.1 INTRODUCTION

Adaptive radiations are characterized by considerable divergence among phenotypic and genetic characters within a group of closely related taxa (Schluter 2000). Furthermore, repeated independent evolution of similar phenotypes in isolated habitats is also a common feature of adaptive radiations (Futuyuma 1998). Such similar (or "parallel") phenotypes may evolve when isolated populations experience similar environments or ecological conditions (Futuyuma 1998; Schluter et al. 2004). Parallel phenotypic evolution provides strong support for the role of natural selection (Schluter and Nagel 1995; Rundle et al. 2000) as closely related taxa share a recent common ancestor, thus they share similar genetic make-up and most likely similar morphologically. These taxa then proceed to evolve in along a similar evolutionary trajectory, although isolated from one another, within similar environments or other evolutionary pressures (Futuyuma 1998). Evolutionary pressures range from strong selection for specialization to exploit scarce resources due to severe competition or defenses against intense predation.

Among fishes, recurrent or parallel evolution is particularly common and definitively observed in guppies (Reznick et al. 1996), cichlids (Ruber et al. 1999; Duponchelle et al. 2008), salmonids (Pigeon et al.1997; Derome et al. 2006; Østbye et al. 2006; Landry et al. 2007), and sticklebacks (Schluter et al. 2004; Colosimo 2004; Boughman et al. 2007). The recurrent evolution of trophic types among fishes has been particularly common (Rüber et al. 2000; Hanel & Sturmbauer 2000), and is thought to result via the repeated invasion of freshwater by marine

ancestral phenotypes (Walker & Bell 2000; Schluter et al. 2004). This has been classically demonstrated in the formation benthic and limnetic forms in coregonids (Bernatchez & Dodson 1990; Pigeon et al. 1997) and sticklebacks (Schluter et al. 2004). Colonization of a new environment frequently leads to the occurrence of divergent morphs, often in sympatry, exhibiting variation in morphological and/or behavioural traits typically associated with foraging (Fraser et al. 1998; Adams & Huntingford 2002a; Klemetsen et al. 2006; Adams et al. 2008). A combination of low species diversity and relative isolation may promote sympatric divergence (Schluter 1996; Adams et al. 2008), and the lack of species diversity coupled with high endemism is likely to have encouraged such patterns within the Malili Lakes on Sulawesi, Indonesia.

Within the adaptively radiating family Telmatherinidae are three major genera *Telmatherina*, *Paratherina*, and *Tominanga* endemic to various regions within the Malili system. The genus *Paratherina* is distributed throughout Lakes Mahalona and Towuti with at least 2 species: *P wolterecki* and *P striata* documented within Lake Mahalona (Kottelat 1990; Herder et al. 2006b). Both species exhibit limited conspicuous morphological differences and species identification keys (Kottel 1990) describe overlap among numerous meristic characters (number of dorsal, pelvic and anal fin rays) rendering species identification very difficult in the field. Only one species of *Telmatherina* is distributed widely throughout the Malili Lakes. Most species of *Telmatherina* are endemic to Lake Matano and can be categorized into three lineages each possessing specialized skull and jaw morphology allowing the exploitation of different resources (Roy et al. 2007). Lake Matano is very limited in productivity (Haffner et al. 2001). Thus, hard selection on jaw morphology is likely the driving mechanism behind the patterns of

divergence in *Telmatherina* (Roy et al. 2007).The *Paratherina* of the similarly oligotrophic Lake Mahalona (Haffner et al 2001; Sabo et al 2008) appear to have evolved in parallel as the distribution of these two species primarily to littoral and pelagic habitats.

In this paper, I present evidence of genetic and ecological divergence in sympatric forms of *Paratherina* from Sulawesi's Malili Lakes region. Secondly, we demonstrate that this divergence closely resembles that of the Lake Matano *Telmatherina* species flock and that the evolution of pelagic and benthic forms evolved independently in the two systems. These data provide further evidence that the primary force driving diversity within this radiation is ecological speciation rather than introgressive hybridization.

6.2 METHODS

Study sites and sample collection

Approximately 100 fish were collected using beach seine and light trap netting from each of three sites in Lake Mahalona in May-June 2005 and June 2006 (Fig. 6.1). All specimens were fin-clipped and collected fin tissue was stored in 95% ethanol. Approximately 30 individuals from each site were euthanized and preserved whole in 95% ethanol.

DNA extraction, microsatellite genotyping & mtDNA sequencing

DNA was extracted from fin clip samples following the plate-based extraction method of Elphinstone et al. (2003) and resuspended in 50 µl of Tris-EDTA buffer (10mM Tris, 1.0mM EDTA, pH 8.0).

An approximately 1206 bp fragment of the mtDNA cytochrome *b* gene was amplified in 24 individuals with forward primer GLUDG-5 (5'-TGACTTGAARAACCACCGTTG-3') (Palumbi 1992) and reverse primer CBtelm-R (5'-GTGGAGGAGGGGTACGACTA-3') (Roy *et al.*, 2007a) using the following PCR conditions: initial denaturation at 95°C for 2 min, followed by 35 cycles of 95°C for 10 s, 60°C for 10 s, 72°C for 10 s, followed by a final extension at 72°C for 1 min. PCR products were then sequenced using ABI Bigdye Terminator v3.1 on an ABI 3130xl Genetic Analyzer.

All sequence-based phylogenetic and molecular evolutionary analyses were conducted using MEGA v4 (Tamura et al., 2007). Sequences were aligned using the CLUSTALW algorithm. Phylogenetic relationships among samples were constructed by neighbour-joining (NJ) using the Kimura-2 parameter and Maximum-Parsimony, with 500 bootstrap replicates and pairwise base deletion.

Two restriction endonucleases *FokI* and *Sau96I* (New England Biolabs, ME) provided diagnostic restriction fragment length polymorphisms of the *cytochrome b* fragment that identified *Paratherina* to the two primary mtDNAclades (M1-M3 and M2). I used the diagnostic RFLP to identify 240 individuals that were not sequenced for *cytochrome b* to determine clade membership for subsequent microsatellite genotyping and analysis.

A total of 288 fish were then genotyped at seven microsatellite markers developed for use in *Telmatherina antoniae* (Tan12, Tan13, Tan14, Tan16, Tan17.1, Tan24, Tan26; Walter et al. 2007). PCR amplifications were: initial denaturation at 94° C for 2 min; then 24 cycles at 94° C for 10 sec, annealing at primer specific temperature for 10 sec and extension at 72° C for 30 sec;

and a final extension at 72° C for 1.5 min (Walter et al. 2007). The size of the PCR products was estimated using a LiCor 4300 DNA Analyzer with GENEIMAGIR 4.05 software (Scanalytics, Inc). I used MSAv4.0 (Dieringer & Schlötterer 2003) to test for genotyping errors, MICROCHECKERv2.2.3 (van Oosterhout 2004) to check for the presence of null alleles, and GENEPOP4.0 (Raymond & Rousset 1995) to test for Hardy-Weinberg and linkage disequilibrium among loci. To illustrate nuclear divergence between the two mtDNA groups, I calculated genetic distance (*D*) and performed a Principal Coordinates Analysis (PCoA) between the mtDNA clades in GENALEX 6.1 using the microsatellite data (Peakall & Smouse 2006). F_{ST} estimates between individuals assigned to clade 1 and clade 2 were calculated in MSA.

Head shape analysis

To measure putative morphological differences between the two primary *Paratherina* groups, the heads of 20 fish from each group were X-rayed using a SDS X-ray imaging system with a PCCR 812 HS OREX digital scanner. Partial warp analysis of head shapes were made using TPSDIG (Rohlf 2006) TPSRELW (Rohlf 2007) based on 11 landmarks identified by Roy et al. (2007).

Stable isotope analysis

Approximately 0.5g of muscle tissue was removed from the left pectoral region of the preserved fish specimens and dried for at least 48 h to remove residual ethanol. To provide a reference, we included 10 specimens of the shrimp *Caridina lanceolata* from L. Mahalona. The samples were freeze-dried and then pulverized in a SPEX CertiPrep 8000-D ball milling unit (SPEX CertiPrep, Metuchen, New Jersey). Inorganic carbon was removed from the samples by acidifying the

pulverized tissue 2x in 50 μg of 20% HCl solution, followed by drying for 24 h (Bunn et al. 1995).

Approximately 1 μg of tissue was weighed and sealed in tin capsules for analysis of $^{15}N/^{14}N$ and $^{13}C/^{12}C$ ratios on a continuous flow isotope ratio mass spectrometer (IRMS; Finnigan MAT Delta^{Plus}, Thermo Finnigan, San Jose, CA, USA) at the Great Lakes Institute for Environmental Research, University of Windsor. By convention, C and N isotopes are expressed in delta notation (δ), the deviation from standards in parts per thousand (‰), according to the following equation: $\delta X = [(R_{sample}/R_{standard}) - 1] \times 10^3$, where X is ^{13}C or ^{15}N and R is the corresponding ratio $^{13}C/^{12}C$ or $^{15}N/^{14}N$. The conventional standard reference materials used were NIST sucrose for CO_2 and NIST ammonia sulfate for N_2. The analytical error for $\delta^{13}C$ and $\delta^{15}N$ analyses were ± 0.2‰, and ± 0.15‰, respectively.

Phylogenetic reconstruction of independent ecological resource exploitation

Approximately 524 bp of the mtDNA *cytochrome b* fragment from *Paratherina* species used above and *Telmatherina* (see Table 6.1 for GenBank accession numbers) was aligned using CLUSTALW. Both a neighbour-joining and maximum-parsimony tree were constructed using pairwise base deletion in MEGA v4. Samples were then assigned to either pelagic or littoral resource exploiting groups on the basis of their stable isotope signatures (see above for *Paratherina*, and Roy et al. 2007 for *Telmatherina*).

6.3 RESULTS

Paratherina *divergence*

The mtDNA *cytochrome b* based phylogenetic tree showed substantial genetic divergence among *Paratherina wolterecki* (M1, M3) and *P striata* (M2) individuals from Lake Mahalona. Estimated sequence divergence (*p-distance*) between M1, M3 and M2 was 3.2%. Further structuring was evident within the *P. wolterecki* clade with 100% bootstrap support with sequence divergence was estimated at 1.5%, possibly indicative of incipient speciation.

Using nuclear microsatellite data, I found support for a *P.wolterecki-P.striata* divergence using PCoA of the genetic distance (*D*) among individuals. Furthermore, significant F_{ST} estimates of 0.15 ($P < 0.001$; Weir & Cockerham 1984) provide additional evidence for strong genetic differentiation among these samples.

No appreciable differences in enrichment for $\delta^{15}N$ were evident among the three groups, indicating that these fishes are likely feeding at the same trophic level. Marked differences in stable isotope signatures for carbon sources were noted among the three genetic groups; M1 and M3 samples again showed greater similarity than with M2. The differences in $\delta^{13}C$ between the M1/M3 complex and M2 are consistent with other studies using $\delta^{13}C$ values demonstrating littoral and pelagic resource partitioning (Paterson et al. 2006; Roy et al. 2007).

No evidence for marked differences in head shape was found using geometric morphometrics was found among the *Paratherina* groups (Fig. 6.4). The results of warps analyses in TPSRELW extracted 69.1% of the variance in head shape among samples in the first two axes; however no clear discrimination between the fish in each group were noted.

Genetic evidence for independent ecological diversification

The neighbour-joining (not shown) and maximum-parsimony trees (Fig 6.5) showed the same resolved the same phylogenetic relationships. Analysis of 524 bp of the *cytochrome b* gene from both *Paratherina* and *Telmatherina* specimens indicate complete isolation of the *Paratherina* and *Telmatherina* complexes with 100% bootstrap support. When applying ecological designations based on stable isotope signatures, it appears that the two primary eco-morphs have evolved independently.

6.4 DISCUSSION

I detected marked genetic divergence within the *Paratherina* genus of fishes from Lake Mahalona with both mtDNA and nuclear markers. Conservative estimates based on a molecular clock for *cytochrome b* assuming 2% sequence divergence per million years, suggest that M1 and M3 group have been isolated from M2 for at least 1.6 my (Brown et al. 1982). Genetic analyses using independent nuclear markers supported the divergence of *Paratherina wolterecki* (M1, M3) and *P. striata* (M2) (Fig. 6.2b). Divergence patterns observed using presumably selectively neutral microsatellite markers reflect the constraint of gene flow and the role of genetic drift in the differentiation among these groups, and may indicate recent evidence of reproductive isolation and potential hybridization.

Ecological divergence among these groups was supported by differences in carbon sources as indicated by stable isotope analyses. The similarities in $\delta^{15}N$ signatures indicate that all three groups of these fishes are feeding at the same relative trophic position, while the differences in

$\delta^{13}C$ between the M1/M3 samples and M2 are consistent with other studies using $\delta^{13}C$ values demonstrating littoral and pelagic resource partitioning (Paterson et al. 2006; Roy et al. 2007).

The occurrence of distinct genetic structuring and the exploitation of littoral and pelagic resources bears a resemblance to patterns observed in the *Telmatherina* species flock endemic to Lake Matano, the headwaters of the Malili system. Differences in jaw morphology, diet, and ecological specialization were noted for three lineages of *Telmatherina* endemic to Lake Matano (Roy et al. 2007). Of the three *Telmatherina* lineages, clades 1 and 3 could be considered pelagic feeders and clade 2 a littoral feeder (Roy 2006). Phylogenetic analysis of the *cytochrome b* gene from both *Paratherina* and *Telmatherina* specimens indicate complete isolation of the *Paratherina* and *Telmatherina* complexes, indicating of independent evolution of the fishes exploiting littoral and pelagic resources in the two lakes. The nutrient-poor characteristics of both Lakes Mahalona and Matano, likely subjected colonizing *telmatherinid* ancestors to similar selective pressures. The independent evolution of these genetic groupings and their respective ecological specializations provide the first example of parallel evolution in the Telmatherinidae radiation.

Concluding remarks

Parallel trends in evolution of phenotypes are believed to be adaptive (Zhang 2006). Here I present evidence of significant genetic structuring using both nuclear and mtDNA markers within the genus *Paratherina* from Lake Mahalona, Sulawesi. The genetic structure is consistent with differences in carbon sources as indicated by stable isotope analyses. This parallelism in the evolution of similar feeding strategies via prey selection is a strong indicator of the role natural

selection independently acting on similar genetic complexes. The distinct ecological divergence between the species in both lakes provides sufficient evidence of similar selective ecological pressures within the Malili system. This paper strengthens support for the primary role of ecological selection in the speciation process (Schluter & Nagel 1995; Futuyuma 1998), particularly within the Telmatherinid radiation (Roy et al. 2007).

6.5 REFERENCES

Adams CE, Huntingford FA 2002. The functional significance of inherited differences in feeding morphology in a sympatric polymorphic population of Arctic charr. *Evol Ecol* 16: 15–25.

Adams E.. Wilson AJ, Ferguson MM 2008. Parallel divergence of sympatric genetic and body size forms of Arctic charr, *Salvelinus alpinus*, from two Scottish lakes *Biol J Linn Soc* 95: 748–757

Bernatchez L & Dodson JJ 1990. Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA analysis. *Evolution* 44:1263–1271.

Boughman JW, Rundle HD, Schluter D. 2005. Parallel evolution of sexual isolation in sticklebacks. *Evolution* 59: 361–373.

Brown WM, Prager EM, Wang A, Wilson, AC 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J Mol Evol* 18: 225-239.

Bunn SE, Loneragan NR, Kempster MA 1995. Effects of acid washing on stable isotope ratios of C and N in Penaeid shrimp and seagrass: Implications for food-web studies using multiple stable isotopes. *Limnol Oceanogr* 40: 622-625.

Cavalli-Sforza LL, Edwards AWF 1967. Phylogenetic analysis: models and estimation procedures. *Am. J Hum Genet* 19: 233-257.

Colosimo PF, Peichel CL, Nereng K *et al.* 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *Plos Biol* 2: 635–641.

Derome N, Duchesne P, Bernatchez L 2006. Parallelism in gene transcription among sympatric lake whitefish *Coregonus clupeaformis* Mitchill ecotypes. *Mol Ecol* 15: 1239–1250.

Dieringer, D. & Schlötterer, C. 2003. Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes* 3: 167-169.

Duponchelle F, Paradis E, Ribbink AJ, Turner GF 2008. Parallel life history evolution in mouthbrooding cichlids from the African Great Lakes. *PNAS* 105: 15475-15480.

Elphinstone MS, Hinten GN, Anderson MJ, Nock CJ 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Mol Ecol Notes* 3: 317–320.

Fraser D, Adams CE, Huntingford FA. 1998. Trophic polymorphism among Arctic charr *Salvelinus alpinus* L., from Loch Ericht, Scotland. *Ecol Freshwater Fish* 7: 184– 191.

Futuyuma DJ 1998. Evolutionary biology 2d ed. Sinauer, Sunderland, Mass. 763 pp.

Haffner GD, Hehanussa PE, Hartoto D 2001. The biology and physical properties of large lakes of Indonesia. In: *The Great Lakes of the World: Food-Web Health and Integrity* eds Manwar M, Heck RE, pp. 183-192. Backuys, Leiden.

Haffner GD, Sabo L, Bramburger A, Hamilton P, Hehanussa P 2006. Limnology and sediment dynamics in the Malili Lakes: What regulates biological production? In: The Ecology and Limnology of the Malili Lakes. Hehanussa PE, Haryani GS and Ridwansyah I, eds. Proceedings International Symposium, Bogor, Indonesia 20-22 March 2006, page 5. LIPI Press, Jakarta, Indonesia. ISBN 979-26-2407-4.

Hanel R, Sturmbauer C.1999. Multiple Recurrent Evolution of Trophic Types in Northeastern Atlantic and Mediterranean Seabreams (Sparidae, Percoidei) *J Mol Evol*50: 276–283

Herder F, Nolte AW, Pfaender J, Schwarzer J, Hadiaty RK, Schliewen U 2006. Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc Roy Soc B* 273: 2209-2217

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.

Klemetsen A, Knutsen R, Primicero R, Amundsen PA. 2006. Divergent genetically based feeding behaviour of 2 sympatric Arctic charr, *Salvelinus alpinus* (L.) morphs. *Ecol Freshwater Fish* 15: 350–355.

Kottelat M 1990. Sailfin silversides (Pisces: Telmatherinidae) of Lakes Towuti, Mahalona and Wawontoa (Sulawesi, Indonesia) with descriptions of two new genera and two new species. *Ichthyol Explor Freshw* 1: 35–54.

Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL 2007. Genetic properties influencing the evolvability of gene expression. *Science* 317: 118–121.

Losos JB, Jackman TR, Larson A, de Queiroz K, Rodriguez-Schettino L 1998. Historical contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279: 2115–2118.

Østbye K, Amundsen P-A, Bernatchez L, Klemetsen A, Knudsen R, Kristoffersen R, Næsje TF, Hindar K. 2006. Parallel evolution of ecomorphological traits in the European whitefish Coregonus lavaretus (L.) species complex during postglacial times. *Mol Ecol* 15: 3983–4001.

Palumbi SR 1992. Marine speciation on a small planet. *Trends Ecol. Evol.*7: 114-118.

Paterson G, Drouillard KG, Haffner GD 2006. Quantifying resource partitioning in centrarchids with stable isotope analysis. *Limnoly Oceano* 51: 1038-1034.

Pigeon D, Chouinard A, Bernatchez L 1997. Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). *Evolution* 51: 196–205.

Raymond M, Rousset F 1995. Genepop Version 1.2: population genetics software for exact tests and ecumenicism. *J Heredity* 86: 248-249.

Richmond JQ, Reeder TW 2002. Evidence for parallel ecological speciation in scincid lizards of the *Eumeces skiltonianus* species group (Squamata: Scincidae). *Evolution* 56: 1498–1513.

Reznick DN, Rodd FH, Cardenas M 1996. Life history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV Parallelism in life-history phenotypes. *Am Nat* 147: 319–338.

Roy D 2006. The evolutionary history and ecology of *Telmatherina* in Lake Matano: An example of adaptive radiation in an ancient lake. PhD Thesis, University of Windsor, Windsor, ON.

Roy D, Paterson G, Hamilton PB, Heath DD, Haffner GD 2007a. Resource-based adaptive divergence in the freshwater fish *Telmatherina* from Lake Matano, Indonesia. *Mol Ecol* 16: 35-48.

Rüber L, Meyer A, Sturmbauer C, Verheyen, E 2001. Population structure in two sympatric species of the Lake Tanganyika cichlid tribe Eretmodini: evidence for introgression. *Mol Ecol* 10: 1207–1225.

Rundle HD, Nagel L, Boughman JW, Schluter D 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* 287: 306-308.

Sabo E, Roy D, Hamilton PB, Heahanussa PE, McNeelyR, Haffner GD 2008. The plankton community of Lake Matano: factors regulating plankton composition and relative abundance in an ancient, tropical lake of Indonesia. *Hydrobiologia* 615:225–235 DOI 10.1007/s10750-008-9560-4

Schluter D 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50:1766–1774.

Schluter D 2000. The ecology of adaptive radiation. Oxford University Press, Oxford.

Schluter D, Nagel LM 1995. Parallel speciation by natural selection. *Am Nat* 146: 292–301.

Tamura K, Dudley J, Nei M, Kumar S 2007 MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4: 535–538.

Walker JA, Bell MA 2000. Net evolutionary trajectories of body shape evolution within a microgeographic radiation of threespine sticklebacks (*Gasterosteus aculeatus*). *J Zool (London)* 252: 293– 302.

Walter RP, Ovenden JR, Heath DD 2007. Characterization of four tetranucleotide and six dinucleotide microsatellite markers for use in *Telmatherina antoniae* and related species. *Mol Ecol Notes* 7: 651-653.

Zhang J 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat Genet* 38: 819-823.

Table 6.1 *Telmatherina* sequences from Roy et al. (2007) accessed via GenBank and used for construction of the *Paratherina-Telmatherina cytochrome b* phylogeny in Figure 4.

Accession no:	Haplotype/Species	Genetic Grouping	Stable Isotope Position
DQ002533	1	Clade 1	Pelagic
DQ002518	14	Clade 1	Pelagic
DQ002519	15	Clade 1	Pelagic
DQ002524	20	Clade 2	Littoral
DQ002525	21	Clade 2	Littoral
DQ002526	22	Clade 2	Littoral
DQ002527	23	Clade 3	Pelagic/Generalist
DQ002530	26	Clade 3	Pelagic/Generalist
NC004385	*Melanotaenia lacustris*	Outgroup	

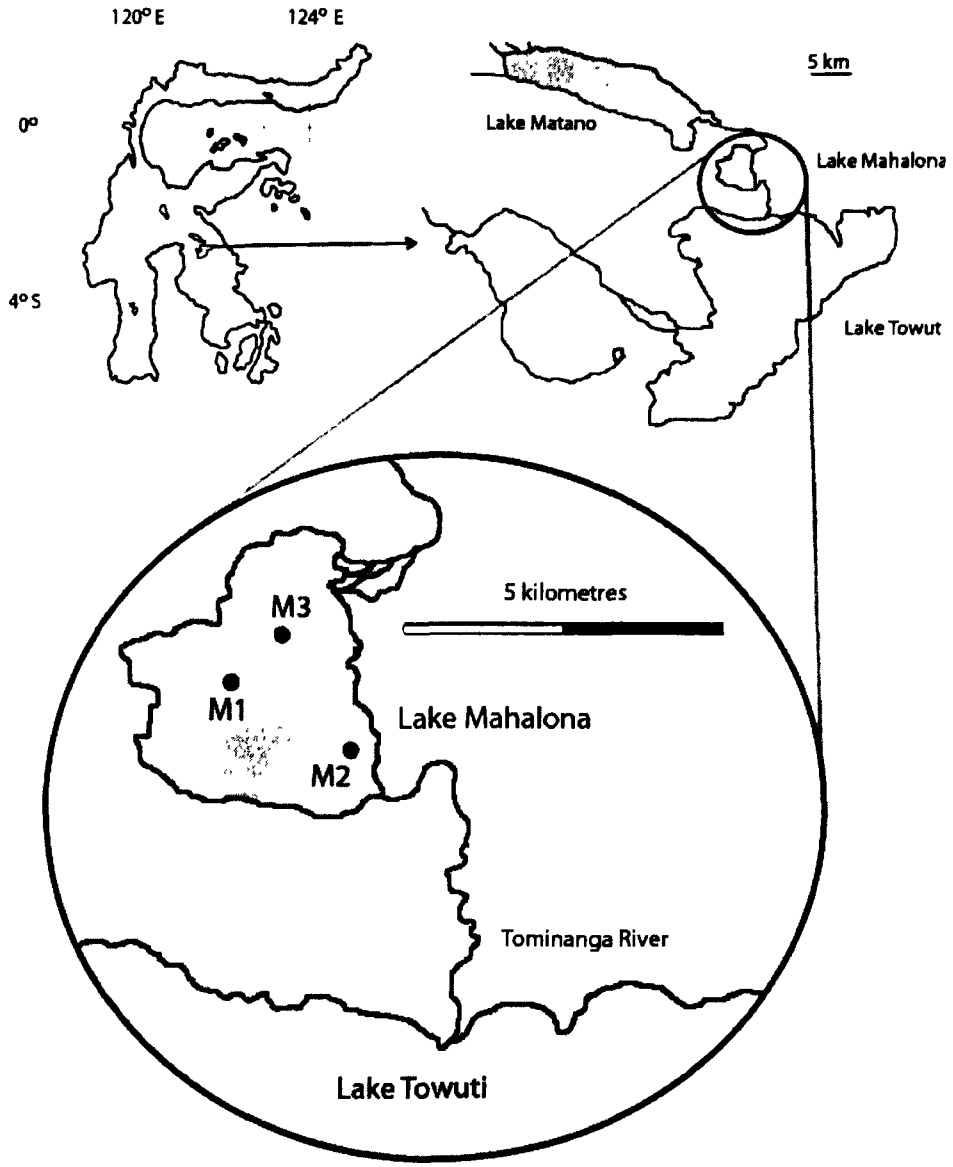


Figure 6.1 — Malili Lakes Region, Sulawesi, Indonesia. M1 and M3 are offshore sampling sites, M2 is a littoral sampling site.

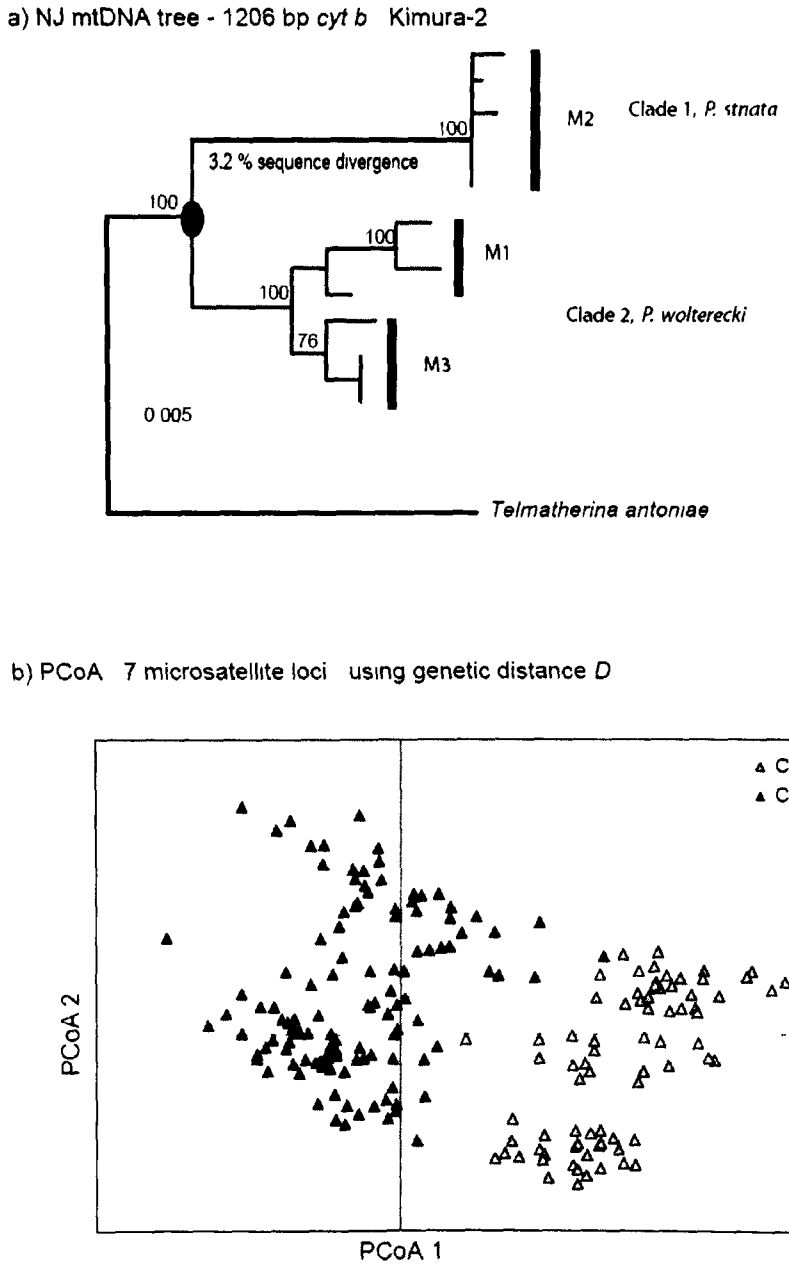


Figure 6.2 — a) Phylogenetic relationship (Neighbour-joining, Kimura-2 parameter) among *Paratherina* using 1206 bp *cytochrome b* mitochondrial sequence. b) Plot of Principal Coordinates Analysis (PCoA) for genetic distance (*D*) among individuals using seven microsatellite loci. Individuals are coded according to their mtDNA clade membership.

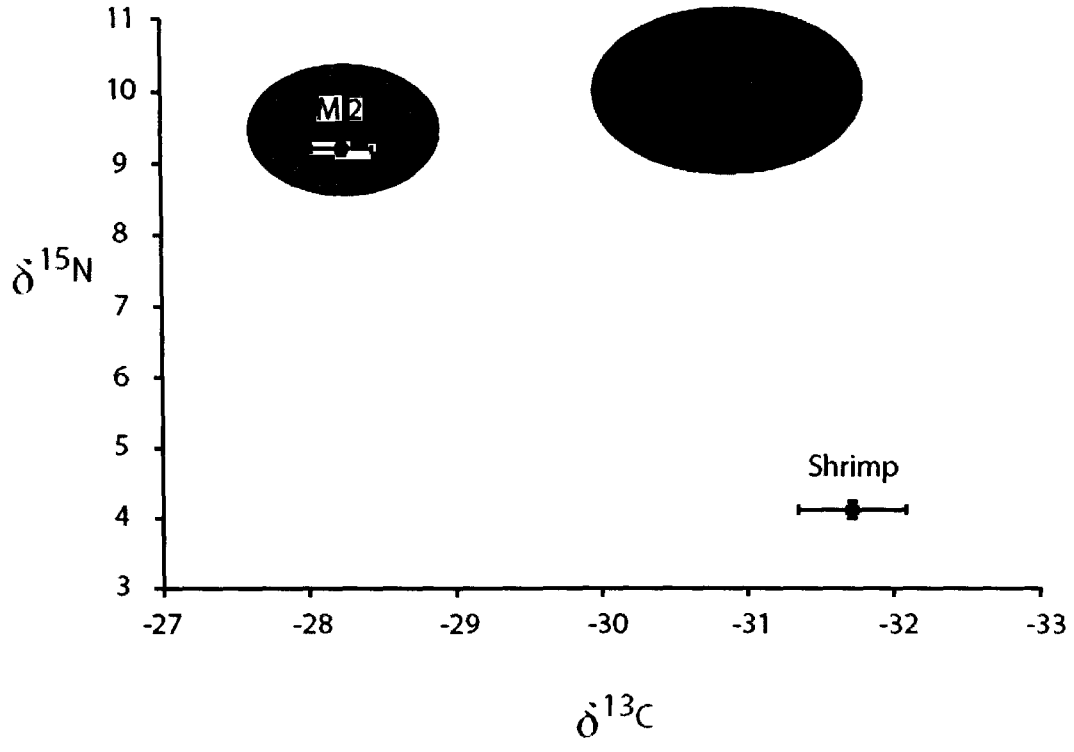
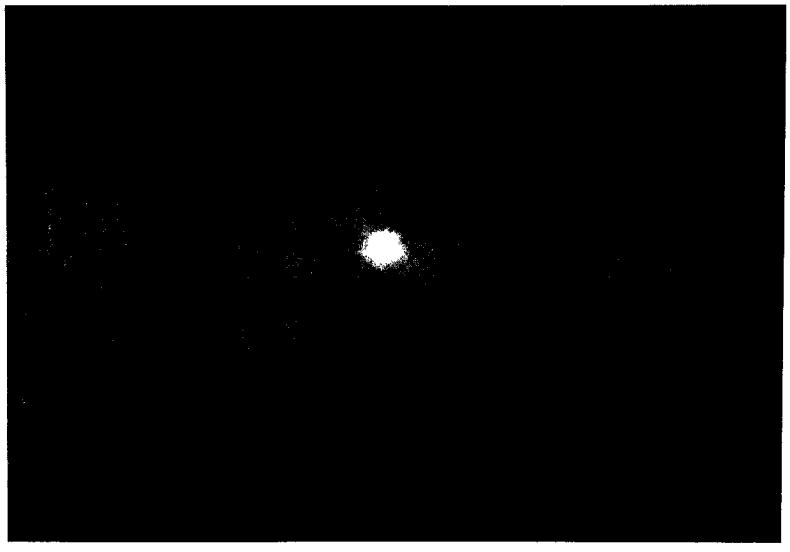


Figure 6.3 — Mean nitrogen ($\delta^{15}N$) and carbon ($\delta^{13}C$) stable isotope signatures for *Paratherina* (M1, M2, M3) and *Caridina* (Shrimp) sampled from Lake Mahalona, error bars indicate standard error. Red and green colours correspond to littoral and pelagic resource signatures.

a) *Paratherina wolterecki*



b) *Paratherina striata*

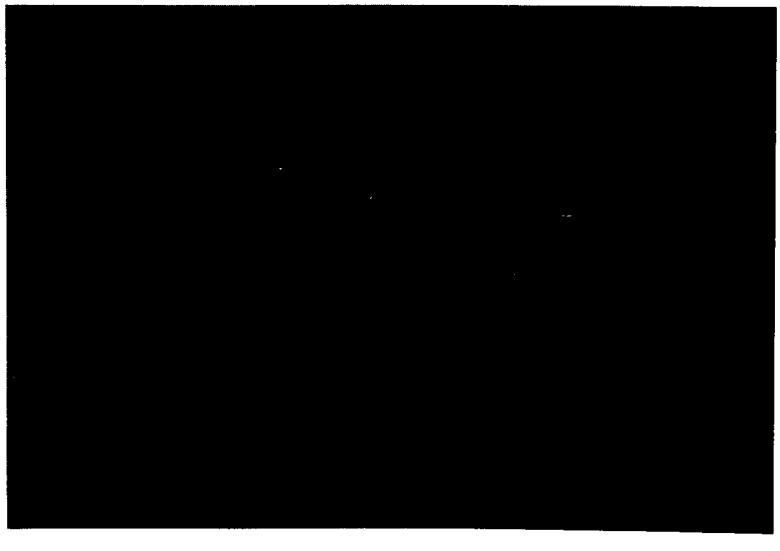


Figure 6.4 — Images of 11 landmark positions (as identified in Roy et al. 2007) used for geometric morphometric analysis of head shape.

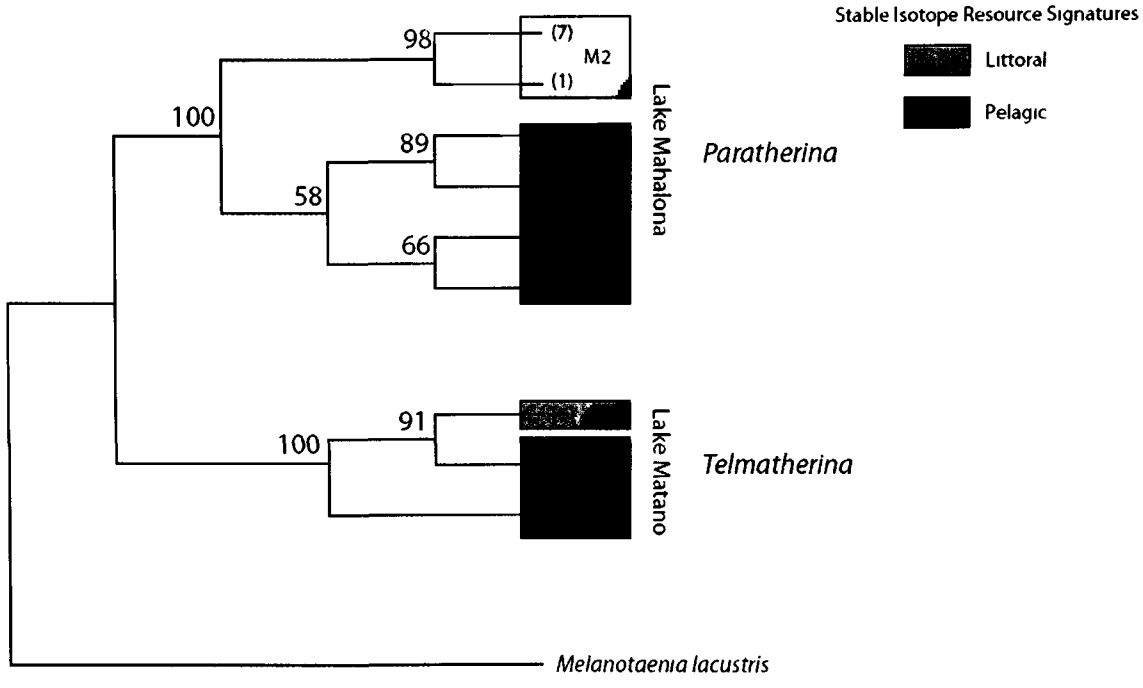


Figure 6.5 — Maximum-parsimony phylogenetic tree for 524 bp of the *cytochrome b* gene from both *Paratherina* from Lake Mahalona and *Telmatherina* from Lake Matano, number of haplotypes in parentheses. Red and green colours correspond to pelagic (red) and littoral (green) stable isotope signatures from as observed in Fig 6.3.

7.0 — General Discussion

Gene flow has classically been thought of as playing a central role in evolutionary processes by homogenizing the genetic composition of populations (Barton and Hewitt 1989; Lenormand & Raymond 2000). The results presented here offers a departure from this classical interpretation by characterizing the differential influences of gene flow (or a lack of gene flow) on the evolutionary history of populations via the dispersal-based strategies of various fish species. Multilocus microsatellite genotyping provided a neutral genetic basis for measuring the effects of drift and gene flow on population genetic structure. For all species studied, I was able to identify genetically distinct groups at the population and species levels; and using genotype assignment techniques, identify dispersing individuals. Low levels of consistent dispersal translated into regular gene flow which served to differentially stabilize Pacific salmonid populations through time. High levels of lake-wide dispersal resulted in admixed populations and sympatric divergence most likely via behavioural and/or ecological specialization in the species *Telmatherina antoniae*. Downstream biases in dispersal resulted in asymmetrical gene flow, ultimately influencing the genetic differentiation of upstream populations in two species of *Paratherina* silversides. Finally, barriers to dispersal make way for parallel ecological specialization, likely prey selection, within the Telmatherinid radiation. This research provides empirical examples of the differential contribution of dispersal to gene flow, population genetic structuring and speciation in fishes. This chapter briefly summarizes the major findings of this thesis providing a context for the study of dispersal and gene flow in fishes and discusses the overall implications for population divergence and speciation.

7.1 CONTEXT

Chapter 1 provided an introduction to the major concepts addressed in this thesis and briefly outlined how these concepts are related to the fundamental theory of evolution. The theoretical basis of population divergence was presented and the role of dispersal was discussed in relation to those mechanisms. Genetic divergence was discussed in relation to the emergence of isolation, local adaptation and speciation. Finally, the selection of fishes as model organisms for studies of dispersal-based evolution was justified.

Throughout this thesis, I rely on genotype assignment methods to detect evidence of both dispersal and gene flow. Ideally, direct measurements are the most desired data to determine contemporary dispersal in an ecological context (Lenormand & Raymond 2000); however, logistical, geographic and temporal scale issues limit direct measurements. While mark-release recapture methods are effective in detecting various aspects of dispersal, they are usually limited to measuring the movement of individuals rather than alleles (Lenormand & Raymond 2000). Therefore, genotype assignment methods are an efficient alternative to direct estimation of migration distances (Smouse & Chevillon 1998; Waser and Strobeck 1998; Lenormand & Raymond; Berry et al. 2004).

7.2 DISPERSAL, GENE FLOW AND TEMPORAL STABILITY IN PACIFIC SALMON POPULATIONS

The stabilizing effects of gene flow for populations through time were demonstrated in Chapter 2. Specifically, population genetic structure (as a result of straying among Chinook salmon) appears consistent with the occurrence of local adaptation. However, genotype assignment

techniques demonstrated that nearly 15% of populations consisted of migrant individuals. Despite such marked gene flow, spatial genetic structure among rivers was preserved—but this gene flow also served to raise effective population sizes, preserving genetic structure within each river through time. These patterns highlight the importance of low levels of consistent gene flow for species where varying degrees of isolation and local adaptation are assumed.

7.3 HIGH DISPERSAL, POPULATION ADMIXTURE AND SYMPATRIC DIVERGENCE

While low levels of gene flow may raise effective population sizes and serve to somewhat homogenize genetic signals, high levels of dispersal can translate into population admixture. The lake-wide dispersal behaviour of Lake Matano's *Telmatherina antoniae* resulted in the sympatric distribution of at least four distinct genetic groups, or populations (Chapter 3). The detection of four populations in admixed proportions throughout this species' native range is indicative of behavioural- and/or ecological-based divergence in sympatry (Chapter 3). These results highlight the fact that dispersal does not equate with gene flow.

Chapter 4 provides a novel contribution to the literature of the Telmatherinid radiation by confirming the presence of a behavioural basis for biased mating patterns within *Telmatherina antoniae*. This chapter uses neutral microsatellite makers to provide evidence for a lack of significant genetic differentiation among colour morphs and shows comparable results to other studies relying on dominant markers (Herder et al. 2008).

7.4 DISPERSAL BIASES AND BARRIERS SHAPE GENETIC DIVERSITY AND ENDEMISM

Chapter 5 provides evidence of genetic connectivity in two endemic *Paratherina* species based on asymmetric riverine gene flow in Lakes Mahalona and Towuti, in Sulawesi, Indonesia. This chapter shows that rivers within the Malili system are not complete isolating barriers as proposed in previous studies (Kottelat 1991; Roy et al.2004), but rather highlights how rivers create non-traditional 'source-sink' patterns across species' distributions. The observed relative genetic isolation of the L. Mahalona *Paratherina* species is of conservation concern in a system undergoing rapid development from mining within the watershed and with substantial changes in habitat quality and structure.

Chapter 6 presents evidence of genetic and ecological divergence in sympatric forms of *Paratherina* from Sulawesi's Malili Lakes and demonstrates that this divergence strongly resembles the resource-based adaptive divergence of Lake Matano's endemic *Telmatherina* species flock. Using mtDNA, I substantiate that the evolution of pelagic and benthic resource exploiting morphs evolved independently in the two systems. This chapter presents data implicating ecological speciation as the principal force driving diversification within the Telmatherinid radiation, in contrast with the work of other researchers (Herder et al. 2006; Schwarzer et al. 2008) in this system.

7.5 FINAL NOTE

While studies frequently highlight a positive relationship between dispersal and gene flow, this thesis serves to emphasize the differential influences of dispersal on genetic structuring and the evolution of natural populations. By providing empirical examples of how dispersal and gene

flow can unite populations, or ultimately serve to drive them apart (ultimately resulting in speciation). This thesis demonstrates the importance of dispersal in the evolution of fishes and provides a better understanding of the role of dispersal in both maintaining and generating new biodiversity.

7.6 REFERENCES

Barton NH, Hewitt GM 1989. Adaptation, speciation and hybrid zones. *Nature* 341, 497–503.

Berry O, Tocher MD, Sarre SD 2004. Can assignment tests measure dispersal? *Mol Ecol* 13, 551-561.

Herder F, Nolte AW, Pfaender J, Schwarzer J, Hadiaty RK, Schliewen U 2006. Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc Roy Soc B* 273, 2209-2217.

Kottelat M 1991. Sailfin silversides Pisces: Telmatherinidae of Lake Matano, Sulawesi, Indonesia, with descriptions of six new species. *Ichthyol. Explor. Fres.* 1, 321–344.

Kottelat M, Whitten AJ, Kartikasari SN, Wirjoatmodjo S 1993. Freshwater Fishes of Western Indonesia and Sulawesi. Jakarta: Periplus.

Lenormand T, Raymond M 2000. Analysis of Clines with Variable Selection and Variable Migration. *Am Nat.* 155, 70-82.

Roy D, Docker MF, Hehanussa P, Heath DD, Haffner GD 2004. Genetic and morphological data supporting the hypothesis of adaptive radiation in the endemic fish of Lake Matano. *J Evol Biol* 17, 1268-1276.

Schwarzer J, Herder F, Misof B, Hadiaty RK, Schliewen UK 2008. Gene flow at the margin of Lake Matano's adaptive sailfin silverside radiation: Telmatherinidae of River Petea in Sulawesi. *Hydrobiologia* 615, 201-213.

Smouse P, Chevillon C 1998. Analytical aspects of population-specific DNA fingerprinting for individuals. *J Hered.* 89, 143–150.

Waser PM, Strobeck, C 1998. Genetic signatures of interpopulation dispersal. *TREE* 13, 43–44.

APPENDIX 1.0 — CHARACTERIZATION OF FOUR TETRANUCLEOTIDE AND SIX DINUCLEOTIDE MICROSATELLITE MARKERS FOR USE IN THE TROPICAL FRESHWATER FISH *TELMATHERINA ANTONIAE* AND RELATED SPECIES

A.1 ABSTRACT

Ten primer pairs were designed from two genomic libraries enriched for $(GACA)_4$ and $(GACA)_7$ in the sailfin silverside *Telmatherina antoniae*. Characterization with 57 *T. antoniae* individuals revealed between three and 30 alleles, with observed and expected heterozygosity values ranging between 0.47 – 0.98 and 0.46 – 0.93, respectively. Eight of the 10 loci conformed to Hardy-Weinberg equilibrium, with no significant linkage disequilibrium detected between loci pairs. These microsatellite markers are intended for use in population genetic studies of *Telmatherina antoniae* and related fishes of the family Telmatherinidae.

A.2. MAIN TEXT

Sailfin silversides (Telmatherinidae) are a family of tropical fishes endemic to the freshwaters of Sulawesi and brackish and mangrove swamps of New Guinea (Saeed and Ivanstoff, 1991; Kottelat, 1993). Some species of *Telmatherina* are ideal for the study of speciation processes as there are numerous adaptively radiated forms (Roy *et al.*, 2004) known from a single lake within the Malili Lakes basin of Sulawesi. The development of highly polymorphic neutral markers is desirable to assess population structure within *Telmatherina antoniae* and may prove useful in related taxa.

Microsatellite enrichment was performed using a modified version of Fischer and Bachman (1998). Genomic DNA was extracted using the Promega Wizard extraction kit (Promega, Madison, USA) from caudal fin tissue of 15 *Telmatherina antoniae*, identified using Kottelat (1991), and supported by RFLP of cytochrome B (Roy *et al.,* in press). Six micrograms of combined genomic DNA was digested with 10 U of RsaI in a 50 μL volume for 2 h at 37° C. Digested product was ligated using the Mlu adaptor complex (21-mer 5' CTCTTGCTTACGCGTGGACTA-3' and a phosphorylated 25-mer: 5' – pTAGTCCACGCGTAAGCAAGAGCACA-3') and 200,000 U T4-DNA ligase in a total volume of 60.0 uL. Two separate microsatellite-containing libraries were constructed using 1 μg of denatured ligated DNA in two separate hybridization reactions using long and short repeat probes, 45° C for $(GACA)_4$ and 51° C for $(GACA)_7$ for 15 minutes. Hybridization to streptavidin-coated magnetic beads was carried out at room temperature on a slowly rotating wheel for 1 h. The beads were washed once with 2x SSC/0.1% SDS for 5 min at 25° C, once with 1x SSC for 5 min at 25° C, and a final wash with 1x SSC for 2 min at 40° C and 46° C, (T_M – 5° C). Enriched DNA was then eluted by washing the beads with 60° C ddH_2O and brought up to a final volume of 50 μL in ddH_2O. Post-enrichment amplification was performed in 50.0 μL PCR using 1 μM 21-mer adaptor primer, 1x PCR buffer (Promega), 200 uM dNTPs and 3 U Taq polymerase (Promega). Amplification was performed at 30 cycles of 94° C for 1 min, 56° C for 45 s, and 72 for 45, with a final extension step performed at 72° C for 5 min. Roughly 100 ng of PCR product was ligated into a pGEM-T vector (Promega) per manufacturer's instructions and the ligation mixture was transformed into competent *Escherichia coli* DH5 cells. White colonies were selected, cultured, and DNA was recovered using thermal lysis (Sambrook *et al.* 1989).

Plasmid DNA was amplified using m13 primers and sequenced using ABI Bigdye Terminator v3.1 on an ABI 3130xl Genetic Analyzer.

Sequence results showed 46 sequences containing novel microsatellite repeats. Twenty-nine primer pairs were designed using Primer3 (Rozen and Skaletsky 2000) and NetPrimer (Premier Biosoft International) software. Allelic variation was evaluated on 57 individuals of *Telmatherina antoniae* from Lake Matano, Sulawesi, Indonesia. The PCR mixture contained approximately 10 ng of template DNA, 1x PCR buffer, 0.4 μM of each primer, 50 μM of each dNTP and 0.5 U Taq polymerase (Sigma) in a total volume of 25 μL. The PCR conditions consisted of an initial denaturation at 94° C for 2 minutes, followed by 35 cycles of 94° C for 10 sec, specific annealing temperature of primer pair for 10 sec (Table A1), 72° C for 30 sec, with a final step at 72° C for 1 min 30 sec. PCR products were separated on 1.8% agarose gels.

Ten of the 29 primer pairs produced clear polymorphic products. Dye-labeled (IR-700, IR-800, MWG-Biotech) forward primers were constructed and used to amplify the 10 polymorphic loci. The size of the PCR products was determined on a LiCor 4300 DNA Analyzer with GeneImagIR 4.05 software (Scanalytics, Inc). The number of alleles per locus ranged from three to 30 and values of observed and expected heterozygosity ranged from 0. 47 – 0.98 and 0.46 – 0.93, respectively (Table A1). Tools for Population Genetic Analysis (TFPGA) v1.3 (Miller 1997, http://www.marksgeneticsoftware.net/tfpga.htm) was used to test Hardy-Weinberg equilibrium with 10000 dememorizations, Bonferroni correction was applied for the number of tests performed.

Eight loci conformed to Hardy-Weinberg equilibrium (HWE; Table A1), while two loci did not conform to HWE due to homozygote excess. No significant ($P > 0.05$) pairwise linkage disequilibrium was detected among the 10 polymorphic loci using Arlequin v3.0 (Excoffier *et al.* 2005). The presence of null alleles was detected using Microchecker v2.2.3 (Oosterhout et al. 2004, http://www.microchecker.hull.ac.uk/) in four loci. Null alleles, population fragmentations or cryptic speciation within *Telmatherina antoniae* may explain the observed departures for loci with significantly deficient heterozygosity.

These primers will be useful in assessment of gene flow and population structure of *Telmatherina antoniae* in Lake Matano, where *T. antoniae* are endemic, and have application for paternity assignment in this and other species. Primer pairs were also effective in amplifying loci in related taxa: including other individuals within the *Telmatherina* species complex: mtDNA clades 2 and 3 (Roy *et al.*, 2007) and the related Telmatherinid *Paratherina wolterecki* (Table 2).

A.3 REFERENCES

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.

Fischer D, Bachman, K (1998) Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.) Biotechniques 24(5): 196-800.

Miller MP (1997) Tools for population genetic analysis (TFPGA). A windows program for analysis of allozyme and molecular genetic data.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Humana Press, Totowa, NJ, pp 365-386

Roy D, Docker MF, Hehanussa P, Heath DD, Haffner GD (2004) Genetic and morphological data supporting the hypothesis of adaptive radiation in the endemic fish of Lake Matano. *J. Evol. Biol.* 17: 1268-1276.

Roy D, Paterson G, Hamilton PB, Heath DD, Haffner GD (2007) Resource-based adaptive divergence in the freshwater fish *Telmatherina* from Lake Matano, Indonesia. Molecular Ecology 16, 35-48.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning. A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York.

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.

Table A1. Characterization of 10 novel microsatellite loci in *Telmatherina antoniae*. Data is based on the amplification of 57 individuals. GenBank = Accession numbers, T_A (°C) = annealing temperature, Range = range in base pairs of the alleles, H_O and H_E indicate observed and expected heterozygosity, respectively, bold heterozygosity values correspond to significant departures ($P < 0.05$) from Hardy-Weinberg equilibrium following Bonferroni correction, * indicates that null alleles may account for heterozygote deficiency

Locus (GenBank)		Primer sequence	Repeat motif	T_A (°C)	$[MgCl_2]$ (mM)	Range (bp)	No of alleles	H_O	H_E
Tan 9	F	ATTGAGCCGTCAGACAAAAA	$(GACA)_5$	52 0	2 5	175-261	25	0 89*	0 92
(DQ914275)	R	AAAGGGGCTAGCTAACTGGA							
Tan 10	F	TAATGAGGGGAGCAGTGGAA	$(GACA)_4$	52 0	2 5	216-226	3	0 51	0 46
(DQ914276)	R	TGTGTGCGTTGGATAAGCAT							
Tan 11	F	CAATCCAGCAGGCTTTTCTT	$(CA)_{17}$	54 1	2 5	113-149	14	0 63	0 67
(DQ914277)	R	CAATGCCTGAGAGGGCTTAG							
Tan 12	F	ACAGATCAAACGGCTCGAAT	$(TG)_{18}$	52 0	2 5	201-229	14	**0.53***	**0.78**
(DQ914278)	R	TGTGCGTGTGAGAGAGTGTG							
Tan 13	F	TATGCTTGGTTGGGTTCACA	$(GT)_{13}$	54 0	2 5	218-228	5	0 65	0 74
(DQ914279)	R	AACCCACCACTGATCAGGAC							
Tan 14	F	GCGAGAACACCGTGTAAATG	$(AC)_{11}$	52 0	2 5	191-197	7	0 47*	0 67
(DQ914280)	R	ACGAGAGTGACGGGGAGTTA							
Tan 16	F	TTGTGGCGACAGACTGATCT	$(AC)_{11}$	57 4	2 5	177-187	3	**0.46***	**0.64**
(DQ914281)	R	AGCTGTGTTAAATGGCAGACC							
Tan 17	F	TTCCTCTCCTCCTGCATGTC	$(GT)_{15}$	54 0	2 5	150-176	19	0 87	0 91
(DQ914282)	R	GAAGGACACGAGGACAGAGG							
Tan 24	F	GTGGTCCACTTTGTGGCTTT	$(GATA)_6N_3(GACA)_5$	52 0	2 5	212-236	7	0 59	0 55
(DQ914283)	R	CTGGTCTTGGTTTGGTATCTATTT							
Tan 26	F	GGGATGCTGAAATGAAATGG	$(AGAT)_{13}$	55 4	2 5	205-273	20	0 98	0 93
(DQ914284)	R	TTGACTTGTCAACCTGACCTTG							

Table A2. Cross-species amplification of 10 primer pairs within the family Telmatherinidae. Samples were screened for PCR amplification of a well-defined band in the expected size range (numbers correspond to the number of individuals amplified, numbers in parentheses are the number of alleles detected) using conditions listed in Table A1 and visualized on 1.8% agarose.

	Telmatherina		*Paratherina*
Locus	mtDNA clade 2	mtDNA clade 3	*P. wolterecki*
Tan 9	3 (4)	3 (4)	13 (14)
Tan 10	10 (5)	6 (4)	11 (3)
Tan 11	11 (4)	6 (2)	18 (9)
Tan 12	13 (9)	6 (9)	18 (13)
Tan 13	13 (4)	6 (2)	18 (6)
Tan 14	13 (3)	6 (4)	18 (5)
Tan 16	13 (5)	6 (6)	13 (8)
Tan 17	10 (9)	5 (6)	18 (12)
Tan 24	13 (5)	5 (4)	18 (10)
Tan 26	13 (9)	6 (10)	18 (15)

VITA AUCTORIS

NAME:	Ryan Patrick Walter
PLACE OF BIRTH:	Rochester, New York, USA
YEAR OF BIRTH:	1978
EDUCATION:	McQuaid Jesuit School, Rochester, NY, USA 1991-1995 Pittsford Mendon High School, Pittsford, NY, USA 1995-1996 State University of New York, College at Brockport, USA 1996-2000, B.Sc. 2000-2002, M.Sc. University of Windsor, Windsor, Ontario. 2004-2009.