2012

Speciation in ancient lakes: Insights from the copepods of Sulawesi

James Joseph Vaillant
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/5587

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
Speciation in ancient lakes: insights from the copepods of Sulawesi

by

James J Vaillant

A Thesis
Submitted to the Faculty of Graduate Studies
through Environmental Science
in Partial Fulfillment of the Requirements for
the Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada

2012

© 2012 James J Vaillant
Speciation in ancient lakes: insights from the copepods of Sulawesi

by

James J Vaillant

APPROVED BY:

______________________________________________
Dr. Trevor E Pitcher
Department of Biological Sciences

______________________________________________
Dr. Daniel D Heath
GLIER

______________________________________________
Dr. Melania E Cristescu, Advisor
GLIER

______________________________________________
Dr. G Douglas Haffner, Advisor
GLIER

______________________________________________
Dr. Lynda D Corkum, Chair of Defense
Department of Biological Sciences

August 16, 2012
I. CO-AUTHORSHIP DECLARATION

I hereby declare that this thesis incorporates material that is the result of joint research undertaken under the supervision of Drs ME Cristescu and GD Haffner (University of Windsor). The material covered in Chapter II is from the article entitled “The Ancient Lakes of Indonesia: Towards Integrated Research on Speciation” which is published in the journal of *Integrative and Comparative Biology*. 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 advice and commentary.

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-authors to include the above material 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 one original paper that has been previously published in peer-reviewed journals, as follows:

<table>
<thead>
<tr>
<th>Thesis Chapter</th>
<th>Publication title/full citation</th>
<th>Publication status</th>
</tr>
</thead>
</table>

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

One of the fundamental questions in biology is the origin of species. Considerable insights into the processes that drive speciation have come from ancient lake systems. In this thesis, I present insights into speciation processes by investigating the species radiations of the ancient Malili Lakes of Sulawesi, Indonesia. The nature of adaptive radiations in the lakes suggests that intraspecific competition for extremely limited resources has driven taxa to adapt to specific habitats and food sources. The copepod populations of Sulawesi reveal that colonization order governs the geographic distribution of zooplankton in freshwater ecosystems. In many Malili Lakes taxa, hybridization between closely related lineages drives diversification, likely by increasing phenotypic diversity within populations. Furthermore, hybridization may be much more common in planktonic taxa than previously thought. Future research in these remarkable habitats is sure to reveal much about the role of hybridization and the origins of biodiversity on earth.
DEDICATION

This thesis is dedicated to my parents, Connie Quinn-Vaillant and Jacques Vaillant, for their unending and unconditional love and support.
ACKNOWLEDGEMENTS

First and foremost, I extend my greatest thanks to my supervisors, Melania Cristescu and Doug Haffner, for their unwavering support and mentorship throughout this journey. Your continued belief in me has fostered both my personal and professional development.

I wish to thank Dan Bock, Sen Xu, Ianina Altshuler, Aibin Zhan, present and former members of the Cristescu lab, Haffner lab, and members of GLIER for their much appreciated comments, conversations, and friendships.

I am especially grateful to Victor Alekseev for his invaluable assistance with the taxonomic work and to Nick Jeffrey and Ryan Gregory for providing genome size estimations. Special thanks to PT Vale Indonesia, Peter Hehanussa, Lili Nuria Lubis, Sinyo Rio, and Pak Dula for providing indispensable logistical support in Indonesia. I also thank my graduate committee, Dr. Daniel Heath, Dr. Trevor Pitcher, and Dr. Lynda Corkum for their support.

Finally, warmest thanks to my parents, Connie Quinn-Vaillant and Jacques Vaillant, for believing in and always being there for me.

This research was funded by the Ontario Graduate Scholarship, National Science and Engineering Research Council Discovery grants to MEC and GDH and by the National Science and Engineering Research Council CREATE training program in Aquatic Ecosystem Health to MEC.
# TABLE OF CONTENTS

DECLARATION OF CO-AUTHORSHIP ................................................................. iii
DECLARATION OF PREVIOUS PUBLICATION .................................................. iv
ABSTRACT ........................................................................................................ v
DEDICATION .................................................................................................... vi
ACKNOWLEDGEMENTS .................................................................................... vii

## CHAPTER

### I. INTRODUCTION

The origin of species .................................................................................. 1
Speciation in ancient lakes ................................................................. 2
References ................................................................................................... 6

### II. THE ANCIENT LAKES OF INDONESIA: TOWARDS INTEGRATED RESEARCH ON SPECIATION

Introduction ............................................................................................... 9
Evolutionary trends in the Malili species-flocks ............................ 13
Speciation patterns in the Malili Lakes .......................................... 20
Towards integrated research on speciation .................................... 24
References ................................................................................................. 29

### III. THE COPEPODS OF SULAWESI REVEAL THE IMPORTANCE OF LOCAL ADAPTATION AND HYBRIDIZATION FOR PLANKTONIC SPECIATION

Introduction ............................................................................................... 38
Materials and Methods ................................................................. 40
Results ......................................................................................................... 44
Discussion ................................................................................................. 48
Conclusion ................................................................................................. 54
References ................................................................................................... 63

### IV. GENERAL DISCUSSION

Final summary ........................................................................................... 71
Closing thoughts ......................................................................................... 72
References ................................................................................................... 77

APPENDIX I .................................................................................................. 79
CHAPTER I
INTRODUCTION

The origin of species

Since the publication of Charles Darwin’s theory of evolution by natural selection in the mid-1800s, the origin of species has been one of the longest standing and most fundamental questions in biology, and yet it remains largely unanswered. For decades biologists have been attempting to elucidate the forces and mechanisms that drive evolution and speciation in order to understand the origins of biodiversity on earth. Traditionally, biologists emphasised the importance of allopatric speciation, the formation of new species as a result of long periods of geographic isolation and subsequent divergence leading to reproductive isolation, regarding it as the primary mechanism for the formation of new species (Coyne and Orr 2004). The biological species concept dominated, and natural selection was thought to be the major driver of evolution (Coyne 1992). However, as speciation research continued, more detailed empirical studies (e.g., African cichlids, threespine sticklebacks, apple maggot fly; reviewed in Via 2001) and theoretical advances (Dieckmann and Doebeli 1999) demonstrated the feasibility and occurrence of sympatric speciation, the divergence and reproductive isolation of populations without geographic isolation. Speciation research moved away from a strictly geographic context and toward the elucidation of underlying processes that drive divergence. It is now understood that speciation is a highly dynamic process that is influenced by many factors including natural selection, genetic drift, genomic rearrangements, and gene flow (Santini et al. 2012). Furthermore, the definition of species has become blurred and numerous species concepts have been proposed which
demark various positions along the species continuum, from genotypic clusters to completely monophyletic clades (de Queiroz 2007). This understanding has raised questions about the relative importance of ecological and genetic drivers of speciation to total biodiversity.

The role of hybridization as a driving force of animal speciation still remains controversial (Seehausen 2004; Mallet 2007; Santini et al. 2012). The traditional view was that gene flow is in direct opposition to divergence by homogenizing genomes between populations through sex and recombination. Thus, hybridization was largely ignored as an important evolutionary process. However, as detailed genetic studies became more feasible and informative there was a resurgence in the exploration of speciation and hybridization (Coyne 1992). By examining multiple genetic markers distributed across the genome, it is possible to reconstruct the evolutionary histories of closely related species and investigate gene flow between diverging populations or emerging species (Butlin 2010). Many recent studies indicate that hybridization between differentiated populations may actually increase phenotypic diversity and, under particular conditions, promote diversification and speciation (Seehausen 2004). Despite these advancements, several concepts remain unclear: 1) the relative contributions of selection, drift, and gene flow toward speciation, 2) the role of hybridization between diverging lineages as a force that increases diversity, and 3) the frequency of hybridization in animals (Schluter 2001; Mallet 2007).

**Speciation in ancient lakes**

Much of the research on speciation processes has been conducted in ancient lakes, which are often considered the aquatic equivalents of islands because of long-term
existence and isolation, discrete geographic boundaries, and unique geological histories (Brooks 1950; Martens 1997). Ancient lakes are extant, long-lived lakes which range from 0.75 to 30 million years in age (Martens 1997; Martens and Schön 1999). These lakes harbour high levels of endemic fauna and are the sites of many rapid, adaptive, and non-adaptive radiation events (Schön and Martens 2004). These lakes have given us many insights into the patterns and processes of speciation. For example, one of the best studied vertebrate radiations is the cichlid fishes of the African Rift Valley lakes (lakes Victoria, Tanganyika, and Malawi). Phylogenetic comparisons of diversification rates and radiation ages in this group have revealed that speciation often occurs rapidly at first, and then slows as lineages age (Seehausen 2006). Whether this slowing is due to a progressive increase in extinction rates or a drop in the speciation rate due to gradual filling of niches remains to be tested, though the latter is more likely in theory (Gavrilets and Vose 2005). It has also been found that certain lineages of cichlids have a higher propensity to speciate than others, as a phylogenetic history of radiation is a strong predictor of subsequent radiations (Seehausen 2006). This indicates that there is an intrinsic component involved in speciation which varies across taxa.

Ancient lakes have also revealed the complimentary roles of environmental fluctuation and introgressive hybridization in driving diversification. Many ancient lakes including Malawi, Tanganyika, Victoria, Caspian, and Titicaca have undergone lake-level fluctuations in response to climatic shifts (reviewed in Cristescu et al. 2010). These fluctuations have certainly influenced the evolution of lacustrine fauna through periods of allopatri, population bottlenecks, and alternating selection regimes. Following lake-level rises, secondary contact and hybridization between previously isolated populations can
then introduce greater phenotypic diversity and promote speciation (Seehausen 2004; Bell and Travis 2005; Koblmuller et al. 2008). In combination, these processes may act as a “species pump” by repeatedly generating novel lineages (Rossiter 1995). Indeed, many ancient lake species-flocks are prone to introgressive hybridization (e.g., Salzburger et al. 2002; Herder et al. 2006a; Marková et al. 2010; Joyce et al. 2011).

The Malili lakes of Sulawesi, Indonesia (lakes Matano, Mahalona, and Towuti) stand out as a particularly attractive natural model system for speciation research as they are the only set of hydrologically connected ancient lakes on earth (Brooks 1950). Sulawesi has the highest percentage of endemism of any island (Bramburger et al. 2008), and its lakes are characterized by unusually low levels of diversity in pelagic phytoplankton, zooplankton, and fish assemblages (Haffner et al. 2001; Roy et al. 2004; Bramburger et al. 2008; Sabo et al. 2008). The lakes also vary in physical dimensions and hydrological connectivity, and feature unique geochemistry and limnological conditions. Thus, the Malili lake system is not only geographically, but also ecologically isolated from surrounding aquatic ecosystems, and provides a singular natural system where speciation mechanisms can be tested. In addition, comparisons with other non-ancient lakes across the island have the potential to provide valuable insights into the insular colonization patterns of various aquatic taxa.

The goal of this thesis is to gain insights into the relative importance of various speciation processes in aquatic systems using the ancient Malili lakes of Sulawesi, Indonesia, as a natural model system. In the second chapter, I introduce the Malili lakes and their unusual physicochemical nature. I then examine emerging evolutionary trends from the findings of recent evolutionary work among various taxa. These trends, which
include adaptive radiation, evolutionary innovation, and hybridization, corroborate the uniqueness of this ancient lake system. I then integrate these findings into a more general framework and propose possible drivers of speciation in the Malili lakes species-flocks. Finally, I outline the overall importance of this study system and its potential utility as a model for future integrated speciation research.

In the third chapter, my main objective is to evaluate the relative importance of various evolutionary forces, including geographic isolation, selection, genetic drift, and hybridization, for speciation in freshwater pelagic invertebrates. To do this, I explore the evolutionary histories of calanoid and cyclopoid copepod populations inhabiting five lakes of Sulawesi, including the Malili lakes, using phylogenetic and population genetic analyses to infer the presence of cryptic species, historical demographic changes, and gene-flow between populations. I find that speciation processes are highly scale dependent, and that hybridization is much more common than previously thought and may play a central role in pelagic speciation. I close the thesis with a summary of my major findings, my thoughts on the causes of high rates of speciation in ancient lakes, and directions for future empirical and theoretical work in this field.
References


CHAPTER II

THE ANCIENT LAKES OF INDONESIA: TOWARDS INTEGRATED RESEARCH ON SPECIATION

Introduction

Ancient lakes as research models

Ranging from 1 to about 30 million years in age, the earth’s oldest extant lakes are regarded as the aquatic equivalents of islands, and typically harbour very high levels of endemic fauna (Brooks 1950; Martens 1997; Cristescu et al. 2010). These ancient lakes have revealed many rapid, adaptive, and nonadaptive radiation events and have provided considerable insights into the major driving forces of speciation; they serve as natural model systems for research into evolution and speciation (Schön and Martens 2004; Seehausen 2006). Their prolonged existence, isolation from ecologically similar habitats, and wide diversity of intralacustrine environments are thought to have accelerated an array of speciation processes (Schön and Martens 2004). Consequently, there is strong empirical evidence for the various theoretical frameworks that have arisen from studies of these systems (Martens 1997; Cristescu et al. 2010). For example, the use of cichlid species-flocks in the East African Rift lakes as model systems has led to advancements in modeling of speciation processes (Danley and Kocher 2001; Gavrilets and Losos 2009). This model system revealed the relative importance of various drivers of adaptive radiation, including ecological opportunity, introgressive hybridization, and sexual selection (Salzburger et al. 2002; Smith et al. 2003; Kocher 2004; Seehausen 2004, 2006; Joyce et al. 2011). Additionally, detailed phylogeographic studies of the cichlid species-flocks have emphasized the importance of habitat fragmentation and
allopatry in generating species diversity (e.g., Rico and Turner 2002; Allender et al. 2003; Genner et al. 2010).

Comparisons among different ancient lakes and their species-flocks have also been useful in identifying specific extrinsic and intrinsic factors responsible for driving radiation events (Fryer 1991; Martens and Schön 1999). Extrinsic factors that are often cited include the long, continuous existence of the lake, topology, depth and fragmentation of littoral habitats, as well as historical fluctuations in climate, temperature regimes, and water levels. Intrinsic factors include specific life-history characteristics such as brooding and philopatry, as well as developmental or genetic constraints such as the modification of the pharyngeal jaw apparatus of cichlid fishes or communication signals in sexually selective species (Fryer 1991; Allender et al. 2003; Schön and Martens 2004). Classical examples of adaptive radiation and speciation in ancient lakes come mainly from studies on the endemic fauna of the African Great Lakes, Lake Baikal, and the Caspian lakes. Many smaller ancient lakes of the world, such as Lakes Titicaca, Biwa, Ohrid, and the Malili Lakes have received relatively little attention. Some of these systems stand out as remarkable settings of adaptive radiation. Our particular interest is the biota of the ancient Malili Lakes of Sulawesi, Indonesia, which is characterized by extremely high levels of endemism and an unusual lack of major taxonomic groups in phytoplanktonic, zooplanktonic, and fish assemblages (Kottelat 1990a; Haffner et al. 2001; Herder et al. 2006b; Bramburger et al. 2008; Sabo et al. 2008). Here, I discuss the unusual biogeochemical nature of these lakes which leaves them not only geographically but also ecologically isolated from surrounding aquatic ecosystems. I present recent findings of evolutionary and ecological studies across various species-flocks, and discuss
emerging evolutionary trends and potential drivers for these speciation processes. Finally, I outline the overall importance of this study system for future research on speciation, especially of planktonic and crustacean taxa.

The ancient Malili Lakes

The Malili Lakes, situated in the central eastern part of the island of Sulawesi, are composed of three major lakes, Matano, Mahalona, and Towuti, and two smaller satellite lakes, Lontoa (or Wawantoa) and Masapi (Fig. 2.1); these represent the only hydrologically connected ancient lake system in the world (Brooks 1950). Lake Matano lies within the Matano fault line, and as such might well be the oldest lake in the system (2–4 million years old; based on fault-line displacements), whereas the other lakes are in areas of complex faulting and currently are estimated to be <1 million years old (based on sedimentary characteristics). Lake Poso is located in a separate drainage system northwest of the central Malili Lakes, and therefore represents a disconnected, yet ecologically and geographically proximal ecosystem. Lake Matano, the best-studied lake of the system, forms the head of the drainage basin. It is a large (164 km²) and exceptionally deep (590 m), steep-sided graben lake (Brooks 1950). The lake is contained in a ferruginous basin with the highest iron concentration of any freshwater lake on earth at 2 mmol l⁻¹ in the epilimnion and as much as 140 mmol l⁻¹ at greater depths (Crowe et al. 2008a, 2008b). The lake is permanently stratified in both temperature and total dissolved solids at about 100 m depth where there is a sharp transition between the warm oxygenated epilimnion, and the cool anoxic metalimnion (Sabo et al. 2008; Crowe et al. 2011). Furthermore, Lake Matano is characterized by low, possibly limiting, levels of dissolved phosphorus (50.2 mmol l⁻¹) and nitrogen (55 mmol l⁻¹) (Sabo et al. 2008), and
potentially toxic concentrations of chromium (180 nmol l\(^{-1}\); Crowe et al. 2008b). These low levels of nutrients and potential long-term toxicity may have considerable implications for the evolution of the biological community of the lake, and it is very likely that the diatom assemblages are controlled by high metal concentrations (PB Hamilton, personal communication). Restricted vertical mixing results in the depletion of nutrients from the epilimnion through sedimentation, and the relatively high concentrations of iron are thought to limit the bioavailability of phosphorus in upper waters (Sabo et al. 2008). Therefore, it is not surprising that Matano has the lowest standing crops of phytoplankton among tropical and temperate ancient lakes with a peak biomass of about 0.013 mg l\(^{-1}\), and harbours a rather simple faunal assemblage (Sabo et al. 2008). Lake Matano discharges into the Petea River, a shallow, fast-flowing river that drops 72 m in elevation before reaching Lake Mahalona. Lake Mahalona is a much smaller (25 km\(^2\)) and shallower (62 m) lake than Matano. It is thought to be more recent in its formation than either Matano or Towuti, although there are no reliable estimates of its age. It is slightly more productive than Matano with respect to supporting subsistence fisheries, and has higher numbers of endemic shrimp and gastropod species, but lower richness of diatom taxa (Brooks 1950; Bramburger et al. 2004; von Rintelen et al. 2011). Mahalona flows out through the Tominanga River, where the water gradually descends 17 m before joining the third major lake of the system, Lake Towuti. Towuti is the largest of the Malili Lakes with an area of 560 km\(^2\) and a maximum depth of 200 m. Although the age of the lake is uncertain, it is estimated to be at least 600,000 years old. It is also remarkably oligotrophic and unproductive (Tierney and Russell 2009); however, it contains the largest number of total species, with at least 13 species of shrimp, 10 species
of molluscs, 10 species of fish, and 154 taxa of diatoms, 39 of which are endemic (Brooks 1950; Kottelat 1990b; Bramburger et al. 2004; Herder et al. 2006b; von Rintelen et al. 2011). Lake Towuti drains through the Larona River and finally discharges at sea into the Gulf of Boni. Nearby, there are also two small satellite lakes, Lontoa (or Wawantoa) and Masapi, which are considerably more isolated and taxonomically impoverished.

Finally, Lake Poso is located about 80 km northwest of the central Malili Lakes in the central metamorphic belt of Sulawesi. It is a large (323 km$^2$; 450m deep), oligomesotrophic lake that lies in a separate drainage basin from the Malili Lakes, and drains northward via the Poso River into the Gulf of Tomini. The lake was likely formed during the collision between the eastern and western plates composing the island, but there is still much uncertainty surrounding the timing of accretion in this region (Audley-Charles 1987; Wilson and Moss 1999). Like the Malili Lakes, the antiquity of Lake Poso is reflected in its biota, as it harbours a number of endemic species-flocks, including aytid shrimps, gastropods, bivalves, and crabs (von Rintelen et al. 2004; von Rintelen and Glaubrecht 2006; von Rintelen et al. 2007; Schubart and Ng 2008; von Rintelen et al. 2011).

**Evolutionary trends in the Malili species-flocks**

*Adaptive radiation*

Adaptive radiations result in the diversification of species through the evolution of morphological or physiological traits that exploit an array of types of resources (Futuyma 1986; Schluter 1996). During rapid diversification, trophic specialization is frequently (although not always) a key feature that characterizes sister species (Fryer
1991; Schönh and Martens 2004). Both theoretical modeling and empirical data indicate that ecologically important morphological features are the first to differentiate during radiation (reviewed by Streelman and Danley 2003; Gavrilets and Losos 2009). This has been observed in the fauna of the Malili Lakes, where nearly all of the species’ radiations described (e.g., telmatherinid fish, gastropods, aytid shrimps, crabs) show trophic specialization, suggesting an exceptionally important role of habitat and resource partitioning in these lakes (von Rintelen et al. 2004; Roy et al. 2007b; Schubart and Ng 2008; von Rintelen et al. 2010).

The best-studied intralacustrine adaptive radiation is that of the endemic sailfin silverside fish *Telmatherina* of Lake Matano. The sailfins comprise two endemic radiations within Lake Matano, the “sharpfins” with seven morphospecies and “roundfins” with three (Kottelat 1991; Herder et al. 2006b; 2008; Pfaender et al. 2010). These species differ in several morphological traits associated with different food sources, including length of the snout, position of the mouth, height of gape, and number of gill rakers (Roy et al. 2004; Pfaender et al. 2010). Further examination of jaw structure and diet (i.e., stomach contents and stable isotope ratios) revealed that each of the morphotypes utilize different feeding strategies that are reflected in their trophic morphologies (Roy et al. 2007b; Pfaender et al. 2010). There are also polymorphisms in coloration of males within each of the lineages, and it is likely that sexual selection is of importance in maintaining these color patterns across varying lake habitats (Gray et al. 2008). However, it is difficult to interpret the historical and contemporary role of color polymorphism in these species as AFLP (Herder et al. 2008), microsatellite (Walter et al. 2009a), and mitochondrial (Roy et al. 2007a) data indicate that color is not correlated
with population structure. Furthermore, different morphospecies show sympatric
distribution at fine scales (Herder et al. 2008), and are likely to disperse around the lake
(Walter et al. 2009b), implicating the importance of habitat isolation through resource
partitioning for the coexistence of these closely related lineages of fish.

Striking cases of adaptive radiation that involve the parallel evolution of adaptive
traits have also been observed in the Malili Lakes. Parallel evolution, or the appearance
of the same trait in independent lineages, is regarded as a signature of natural selection,
as genetic drift is unlikely to produce environmentally correlated changes multiple times
(Schluter and Nagel 1995; Rundle et al. 2000). The observation of this phenomenon is
usually an indication of adaptation to similar environmental pressures. For example, each
of the Malili Lakes and Lake Poso harbours endemic crab species characterized by three
different ecotypes: a generalist, a molluscivore, and a detritivore (Schubart and Ng 2008;
von Rintelen et al. 2011). While representatives from each of these types are always
present in any given lake, examination of phylogenetic relationships between lake
populations suggests that these ecotypes evolved independently at least twice (Schubart
and Ng 2008; Schubart et al. 2008), thereby lending support to the primary role of natural
selection and specialization in feeding preferences.

One of the most notable cases of parallelism is that of the remarkable variation in
trophic specialization in the endemic freshwater gastropods of the genus *Tylomelania*.
While the morphology of gastropods’ radula is generally considered a conserved
character at the species level, nine different morphologies were found in the *Tylomelania*
complex of the Malili Lakes, with nearly identical types occurring in different
populations from different lakes. Furthermore, the molecular phylogeny reveals that in
three of the four major lineages, these characters evolved in parallel following coloniza-
tion of the lakes. This strongly suggests functionality of these traits and a principal role of ecological divergence in the intralacustrine speciation process (von Rintelen et al. 2004; von Rintelen et al. 2011).

Pronounced trophic specialization has also been found in the aytid shrimps of Lake Poso and the Malili Lakes. This is the largest known radiation in the genus *Caridina*, with 14 different species characterized by seven ecotypes (von Rintelen and Cai 2009). Within this species complex, morphological features likely associated with substrate preference were found to be very strong predictors of habitat type, and several of these specializations have evolved more than once in different lineages (von Rintelen et al. 2007, 2010). With most of these species endemic to single lakes, habitat specialization has been suggested as the primary driver of the radiation (von Rintelen and Cai 2009).

Several authors have proposed that the initial stage of the radiation process in ancient lakes is driven by divergence in habitat preference, closely followed by morphological differentiation attributed to trophic specialization (Danley and Kocher 2001; Streelman and Danley 2003; Kocher 2004). The Malili Lakes represent an ideal system to test this model, as the members of a given species-flock are frequently found occupying distinct microhabitats and display adaptive trophic morphologies, and several radiations exhibit the characteristics of early speciation (e.g., shallow divergence between sister species, lack of intraspecific genetic structure, relatively small number of species within the flocks). This constant theme undoubtedly points toward selection due to limited resource availability as the principal driver of divergence.
Evolutionary novelty

Considering the prolonged isolation of ancient lakes, it is not surprising that endemic species possess highly specialized or distinctive characters. The Malili Lakes are exceptional in that they harbour several species that display novel morphological and behavioral traits that are unique to their taxonomic grouping. Here, the species-flock of the limpet snail *Protancylus*, endemic to Lake Poso and the Malili Lakes, exhibits a novel type of brood care. This behavior is not found in any other known limnetic pulmonate snails and consists of ovipositing eggs on the hard shell as a substrate, followed by development of the embryos under the mantle of the female’s shell (Albrecht and Glaubrecht 2006). Interestingly, the use of the shell as a substrate for oviposition is also seen in the pulmonate snails of the ancient Lake Baikal, suggesting that similar selective pressures may have played a role in the development of this behavior (Albrecht and Glaubrecht 2006).

Lake Poso and the Malili Lakes harbour several corbiculid bivalve species. One species, endemic to Lake Poso, is the only sessile (i.e., cementing) species of the lakes and was long thought to have evolved independently. However, a molecular phylogeny reveals that this lineage is very closely related to one of the nonsessile lineages of the lake, and is quite divergent from the other lake species, suggesting rapid evolution of this behavior within its lineage and highlighting the importance of developmental plasticity in adaptation (von Rintelen and Glaubrecht 2006). Furthermore, the molluscivorous crabs of the Malili Lakes possess large molariform chelar teeth, a morphological characteristic known only from marine crabs and, interestingly, those of the ancient Lake Tanganyika (Schubart and Ng 2008). The crabs are closely related to other species from Sulawesi,
making this morphological feature unique within the family Gecarcinucidae. Overall, these innovations and their similarity to marine and other ancient-lake systems suggest that these lakes maintain certain characteristics, such as long-term stability or unusual selection regimes, that promote differentiation, although it is still unclear which forces are responsible for the appearance of such a high degree of specialization.

**Hybridization and speciation**

The role of introgressive hybridization as an important evolutionary force for speciation in animals has been debated for many years (Barton 2001; Mallet 2007). Traditionally, hybridization between distinct populations has been regarded as either a transient process of little evolutionary significance as populations diverge, or in direct opposition to divergence through the continuous exchanging of genes (Barton 2001). More recently, however, hybridization has been gaining support as a process that actually promotes speciation, primarily because it can result in increased variation in quantitative or mating traits, genome rearrangements, or transgressive segregation, with the appearance of “extreme” phenotypes (Rieseberg et al. 1999; Seehausen 2004; Mallet 2007). Indeed, this concept has been supported by studies of several species-flocks from ancient lakes (e.g., Baikalian sculpins, African cichlids) (Salzburger et al. 2002; Kontula et al. 2003; Smith et al. 2003; Seehausen 2004, 2006), including the *Telmatherina* of the Malili Lakes (Herder et al. 2006a; Herder and Schliewen 2010).

Speciation with introgressive hybridization has been demonstrated in the atheriniform fishes of the Malili Lakes, and represents one of the most convincing cases of hybridization in a radiation event. Within Lake Matano, there are two major morphological telmatherinid species-flocks, the “roundfins” and the “sharpfins”, while
the surrounding streams harbour another species. Nuclear AFLPs support monophyly of the “roundfins” and “sharpfins” (Herder et al. 2006a). However, the “sharpfins” possess two mitochondrial lineages. One lineage is sister to the “roundfins”, corroborating common ancestry of both Matano species-flocks, while the other lineage is closely related to the Telmatherina from streams, indicating mitochondrial introgression from the stream telmatherinids into the lacustrine “sharpfin” population (Herder et al. 2006a; Schwarzer et al. 2008). This introgressed population also exhibits substantially higher morphological diversity, suggesting that hybridization may facilitate the evolution of novel phenotypes (Herder and Schliwen 2010). Additionally, at least three other major hybridization events have been shown between other lake and stream Telmatherina populations of the Malili Lakes, leaving open the possibility that reticulate speciation has had a significant role in their diversification as well (Herder et al. 2006a). Hybridization is also highly suspected in the Tylomelania flock of the Malili Lakes, as there is little species-level resolution in their mitochondrial phylogeny and several major morphological features (e.g., shell ribbing) conflict with their placement within mitochondrial phylogenies (Glaubrecht and von Rintelen 2008). Further investigation, including the use of nuclear markers, is needed to confirm this.

As multilocus genetics studies become more widely available, hybridization continues to emerge as an important driver of speciation in many animal taxa (Dowling and Secor 1997; Seehausen 2004). Although it appears that reticulate speciation is more common in fish, this is likely a biased conclusion as very few invertebrate taxa have received the same level of attention. It remains possible that hybridization between moderately divergent lineages is a major creative force in the evolutionary process and is
far more common in nature than traditionally considered. Additionally, while it has been suggested that the same habitats promote both hybridization and adaptive radiation (i.e., novel or perturbed environments) (Seehausen 2004), the specific conditions under which hybridization is likely to generate new species are still unclear. More studies across many different taxonomic groupings are still necessary to better understand these processes.

**Speciation patterns in the Malili Lakes**

Overall, the common themes from each of these studies, including low taxonomic richness, high endemism, and restricted geographic distributions of species within the Malili Lakes, suggest that each of these lakes is characterized by unique and intense physical, chemical, and ecological selection regimes (see Appendix I). Furthermore, each of the species-flocks examined thus far display signatures of adaptive radiation, especially in trophic morphologies, implying that the system is regulated from the bottom up, and that availability of resources is the critical factor governing speciation. Many authors have commented on the simple community structures of the Malili lakes (Haffner et al. 2001; Roy et al. 2004; Bramburger et al. 2008; Sabo et al. 2008). Within the planktonic community, there is a notable lack of entire classes of phytoplankton and centric diatoms, which are typically found in other taxonomically impoverished ancient lakes (Bramburger et al. 2008; Sabo et al. 2008). The zooplankton assemblages of the Malili Lakes are also very simple, composed of single calanoid and cyclopoid lineages, and at least in Matano, a single reported rotifer (Sabo et al. 2008; personal observation). Furthermore, the radiation of telmatherinids into fifteen different species within several intermediate-sized lakes and a number of river and stream ecosystems that present a wide range of heterogeneous habitats can also be considered limited when compared to the
extensive radiations of fish in other ancient lakes, including the 33 species of Baikalian
cottoids (Kontula et al. 2003) and hundreds of species of African cichlids within single
lakes (Seehausen 2006). The absence of cosmopolitan species and simple community
structures characterizing the Malili Lakes suggests that there is a low susceptibility to
new colonizations, perhaps due to limited availability of niche space for new migrants
and the peculiar geochemistry that involves potential chromium-induced toxicity.

Endemism is clearly one of the defining features of the Malili Lakes. Much of this
endemism may be attributed to Sulawesi’s isolation and prolonged existence within
Wallacea. This explains the absence of “primary” freshwater fishes as well as
cosmopolitan phytoplankton and zooplankton species (e.g., centric diatoms, cladocerans)
(Bramburger et al. 2004, 2008; Sabo et al. 2008). Indeed, almost all members of
zooplanktonic, molluscan, shrimp, crab, and fish species-flocks are endemic to the Malili
Lakes and tributaries (Brooks 1950; Roy et al. 2004; Herder et al. 2006b; von Rintelen et
al. 2007; Schubart and Ng 2008). This high level of endemism supports long-term
geographic and ecological isolation of the lakes and limited potential for new
colonizations (Roy et al. 2004). While the Malili Lakes are considered a hotspot for
biodiversity, it appears that the majority of this diversity is contained within a handful of
endemic species-flocks which have radiated within the lakes.

Many species involved in radiation in the Malili Lakes are confined to single
lakes. In the phytoplanktonic assemblage, 32% of the diatom flora is endemic to single
lakes (Bramburger et al. 2004). New evidence shows that a major genetic split exists for
the calanoid fauna between lakes Mahalona and Towuti (see Chapter III). This is
especially surprising, as planktonic species are expected to disperse passively throughout
the lake system. The restricted distributions of some species also indicate that despite
being interconnected, each lake contains distinct biotic communities. Lake Matano
harbours ten species of *Telmatherina*, while Mahalona and Towuti share at least five
different species from *Telmatherina* and the closely related genera *Tominanga* and
*Paratherina* (Herder et al. 2006a). Although there is some support for gene flow between
closely related species of lakes Matano and Mahalona through introgression and
backcrossing with riverine populations (Schwarzer et al. 2008), the lacustrine
telmatherinids of Lake Matano are not found in the Petea River, which links the two
lakes. The Petea River also marks a distinct distributional barrier for freshwater crabs. Of
the two independent colonizations by crabs in the Malili Lakes, one lineage has radiated
into two distinct genetic clades, one shared between Mahalona and Towuti, and the other
harboured solely by Matano (Schubart et al. 2008; Schubart and Ng 2008). Once again, a
similar pattern is found in the shrimps. The Malili Lakes have been colonized twice by
representatives of the genus *Caridina* (von Rintelen et al. 2010). One colonization
resulted in the radiation of *C. lanceolata*, in which a major genetic divergence exists
between the population of Matano and that of Mahalona/Towuti (Roy et al. 2006; von
Rintelen et al. 2010) and the second colonization resulted in a larger radiation of 13
species, 11 of which are present in only one or two lakes (von Rintelen et al. 2010). There
have also been three independent colonizations of the Malili Lakes by members of the
gastropod genus *Tylomelania*. Two of these colonizing lineages are still primarily
confined to lakes Mahalona and Towuti. The third contains species distributed throughout
the Malili Lakes, with several restricted to single lakes, mainly Matano (von Rintelen et
al. 2004). The interconnectedness of the Malili Lakes suggests that most taxa should be
able to disperse both upstream and downstream throughout the system (Brooks 1950). The restriction of so many species to single lakes, despite hydrological connectivity, is curious and suggests the prevalence of ecological (biotic or abiotic) obstacles to contemporary dispersal and colonization.

Within the diatom assemblage, taxonomic richness at the generic level varies greatly among lakes, suggesting differential success of various species in each lake. The ultra-oligotrophy and unique physical, chemical, and limnological conditions of each of these lakes are most likely responsible for limiting the success of different taxa of diatoms (Bramburger et al. 2008; Sabo et al. 2008; see Appendix I). The specialized feeding morphologies of the gastropod and shrimp species-flocks likely reflect this scarcity in phytoplankton and periphyton crops (von Rintelen et al. 2004; von Rintelen et al. 2010). The telmatherinid fish of Lake Matano also display variation in feeding morphologies and preferences indicative of resource partitioning (Roy et al. 2007b; Pfaender et al. 2010). Hence, it is likely that such extreme resource limitation has strongly promoted specialization at each trophic level, and speciation based on resource partitioning has reduced intraspecific competition for resources. Furthermore, local adaptation of resident species may enhance priority effects and reduce the success of establishment of new colonists (De Meester et al. 2002). This effect would be even more pronounced in the resource-limited Malili Lakes, as locally adapted populations would retain an exceptional advantage over migrants from neighboring lakes; this would explain the limited geographic distributions of many species to single lakes and the absence of perhaps less specialized cosmopolitan species. Thus, I conclude that intralacustrine radiations within the Malili Lakes are primarily driven by the unique physical, chemical,
and limnological conditions of each lake which limit productivity, while competition and specialization for lake-specific food sources are responsible for the lack of interlacustrine dispersal and colonization of many species. These species are reproductively isolated through habitat isolation, as they are well adapted to their own habitats, but maladapted to those of sister lineages.

Alternatively, it is possible that geography and allopatric processes are also partly responsible for the distributions of species-flocks in the Malili Lakes. There is some evidence of fluctuation in water level from acoustic and seismic surveys (JM Russell, unpublished data), such as submerged river valleys, in lakes Matano and Towuti that may have caused changes in habitat or periods of isolation between different lakes. It is likely that these periods of isolation may have generated diversity among lakes and facilitated local adaptation within lakes. This corroborates many of the large genetic splits between species of different lakes. However, these processes do not explain the maintenance of restricted distributions following rejoining of the lakes or the presence of multiple intralacustrine radiations with parallel evolution of adaptive traits. While it remains likely that the complex radiations of species-flocks in this system are the result of sympatric and parapatric processes, the potential role of allopatry cannot be discounted. Such stochastic natural events, including sudden geological or limnological changes and recent extinctions, may have played an important role in determining the community composition and the distribution of species-flocks in the lakes (Whitten et al. 1987).

**Towards integrated research on speciation**

Overall, when the biotic and abiotic characteristics of the Malili Lakes are considered, we are presented with a unique opportunity to study speciation processes.
The low taxonomic richness and simple community structures within these lakes facilitate the study of the specific forces driving radiations in the system by limiting the total number of abiotic and biotic interactions to consider. In such a system, a comprehensive characterization of the physicochemical conditions and the entire food web of the lakes will allow for the quantification of physical and biological interactions. Moreover, ecological-genomics studies in the system can link the direct genomic effects of environmental factors (e.g., metal toxicity, nutrient availability) to genotypes and infer population-level responses and speciation events. Detailed characterization of biotic interactions, such as bioenergetics and food-web dynamics, may then lead to the prediction of community-level changes. Furthermore, the interconnectedness of the lakes offers an ideal arena to observe passive dispersal and speciation patterns of planktonic species in the differing environments of each lake. The inclusion of Lake Poso as a proximal but segregated system would also reveal much about the role of geography and dispersal ability in the evolution of various study taxa. As more advanced geological studies are conducted on the lake basins, it will become possible to add timelines to various radiations. This may include correlating physical conditions of the lakes with the occurrence of speciation events, the age and rate of radiations events, and using fossils to calibrate molecular clocks in many taxa. It appears that the radiations of shrimp, fish and possibly gastropods may be in their early stages, as there is a lack of genetic structure at the intra-specific level and divergences of mitochondrial sequences between sister species are still very small (von Rintelen et al. 2007; von Rintelen et al. 2010; Herder and Schliewen 2010). Indeed, the utility of this system for studying radiations has already been highlighted by several authors, as the small size and intermediate complexity of the
radiations opens a rare opportunity to observe and follow the early stages of a radiation and investigate the principal drivers of this process (Herder et al. 2006a; Herder and Schliewen 2010).

Perhaps even more noteworthy is the potential utility of the Malili Lakes for studying crustacean radiations, particularly because of their young ages and intermediate sizes. While many other remarkable crustacean radiations have occurred within ancient lakes, including the amphipods of Lakes Ohrid, Titicaca, and Baikal (in which there are over 300 species) (Kamaltynov 1992), over 200 species of ostracods in both Lake Tanganyika and Baikal of which 90% are endemic (Martens 1994), many of these radiations are old and extensive. Although large-scale comparisons between these radiations have been helpful in revealing the relative importance of certain factors responsible for speciation (e.g., brooding, trophic specialization, sexual selection), the extremely old age of these radiations and the large numbers of species have proven difficult in demonstrating common drivers of diversification (Martens and Schön 1999). While the Malili Lakes still possess very high crustacean endemicity, the number of species within these young radiations is relatively small. Thus, they may provide a key step in linking crustacean diversity to specific evolutionary forces.

The Indonesian ancient lakes represent a remarkable system for evolutionary study. Their unusual physical conditions and varying degrees of isolation and interconnectedness promotes them as an ideal natural model for research on speciation. The use of an integrated approach involving geological, limnological, ecological, and genomics will certainly provide valuable insights into the mechanisms of speciation in aquatic taxa. Important conservation issues can also be addressed as increasing
anthropogenic activity surrounding the lakes threatens endemic fauna (Whitten et al. 1987; Sabo et al. 2008). The Malili Lakes will unquestionably continue to reveal much about radiation and may potentially lead to comprehensive understanding of speciation processes in natural systems.
Fig. 2.1: The ancient Malili Lakes and nearby Lake Poso of Sulawesi, Indonesia. Dashed arrows represent the direction of water flow. Shaded regions indicate water depth in meters.
## References


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


CHAPTER III
THE COPEPODS OF SULAWESI REVEAL THE IMPORTANCE OF LOCAL ADAPTATION AND HYBRIDIZATION FOR PLANKTONIC SPECIATION

Introduction

A fundamental goal of biology is to understand the forces that create and maintain diversity. Traditionally, hybridization between closely related species was viewed in direct opposition to diversification, and thus was considered to be of little evolutionary importance (Mayr 1963). Many recent studies, however, have found that hybridization increases genetic diversity and can result in the acceleration of evolutionary processes such as adaptation and speciation by increasing phenotypic variation and the potential for selection (Seehausen 2004; Mallet 2007). While hybridization is quickly gaining acceptance as an important driver of speciation, its occurrence and contribution to biodiversity remains unclear (Mallet 2007; Santini et al. 2012).

Important insights about speciation processes and the role of hybridization hybridization in aquatic systems have come from studies in ancient lakes (> 1 million years old), where adaptive radiations have occurred in a wide range of taxa (Martens 1997; Cristescu et al. 2010). Here, species are often found to diverge rapidly into species-flocks of closely related sister species. In many cases, environmental fluctuations, especially water level changes, are considered to be responsible for creating periods of allopatry and subsequent divergence between populations (Cristescu et al. 2003; Genner et al. 2010). In several cases, introgressive hybridization has been implicated as an important feature of these adaptive radiations by increasing phenotypic diversity (Herder et al. 2006; Stelkens et al. 2009) and generating new species (e.g., Smith et al. 2003).
Combined, these evolutionary forces have the potential to drive rapid speciation and adaptive radiations.

Despite major progress in the understanding of speciation processes from studies in ancient lakes, much of the focus has remained on benthic species. Comparatively little is known of speciation processes in planktonic species because the common ecological barriers involved in other adaptive radiations, like habitat isolation, are less conspicuous in the open-water realm. Furthermore, many planktonic species display life-history characteristics thought to oppose diversification, like large population sizes and high dispersal rates. However, there is growing evidence that planktonic taxa are very diverse and often constitute cryptic, morphologically indistinguishable, species with distinct allopatric distributions (Goetze 2003; Adamowicz et al. 2007).

The ancient Malili Lakes of Sulawesi, Indonesia, are centered in the unique biogeographic region of Wallacea. They are the only known interconnected ancient lake-chain on Earth and thus provide a unique opportunity to compare the roles of geography and dispersal in planktonic speciation. They have been the site of many adaptive radiations including telmatherinid fishes, gastropods, shrimps, and crabs (reviewed in Vaillant et al. 2011; von Rintelen 2011). In all cases, these species-flocks display adaptations associated with feeding morphology or habitat preference. While geographic proximity and hydrological connectivity allows dispersal between the lakes, many of these species remain restricted to a single lake, suggesting that selection governs species distributions in the system (Vaillant et al. 2011). Furthermore, hybridization has been found to occur between lake and stream telmatherinid fish species (Herder et al. 2006), and is suggested in the gastropod species-flock (Glaubrecht and von Rintelen 2008). In
this study, I compare the relative importance of geographic isolation, hybridization, and ecology for speciation in planktonic taxa. I use cytogenetic, phylogenetic and population genetic analyses of the diaptomid (copepoda: calanoida) and cyclopoid (copepoda: cyclopoida) populations of Sulawesi at two spatial scales. At a large scale, I compare three geographically isolated lake systems of Sulawesi (Lake Tondano, Lake Poso, and the ancient Malili lakes) to evaluate the role of long distance dispersal and colonization. At a local scale, I investigate population structure and possible cryptic species in the interconnected Malili Lakes (Matano, Mahalona, and Towuti).

**Materials and Methods**

*Sample collection*

I surveyed five lakes on the island of Sulawesi, Indonesia: Lakes Tondano, Poso, and the three lakes of the Malili lake system (Matano, Mahalona, and Towuti). Three to five sites were sampled per lake for a total of 19 sites (Table 3.1; Fig. 3.1a). Zooplankton was sampled through 10 m vertical tows using a 62-μm-mesh net (1-m diameter). Specimens were immediately preserved in 95% ethanol. Taxonomic identification was done using the key of Reddy (1994).

*Genome size estimation*

I estimated genome sizes of four individuals from each diaptomid population to investigate polyteny or polyploidy as a source of cryptic speciation using the Feulgen image analysis densitometry method as described by Hardie et al. (2002). Genome sizes were estimated using a minimum of 20 and maximum of 50 nuclei. Optical densities were converted to picograms using chicken (*Gallus gallus domesticus*) blood as a standard. Means and standard errors were calculated from four individuals of each lake.
A one-way ANOVA was performed with the STATISTICA v. 8.0 software package to test for differences between all lake populations (StatSoft Inc. 2007).

**DNA extraction, amplification, and sequencing**

Total genomic DNA (gDNA) was extracted from single individuals using a modified proteinase K method (Schwentk et al. 1998). For calanoid copepods, a total of 418 individuals were analyzed for the mitochondrial cytochrome c oxidase subunit I (COI) gene (Table 3.1). A subset of these individuals, distributed within each of the major COI clades, were analyzed for the nuclear internal transcribed spacer 1 (ITS1) region ($N = 49$), and 18S and 28S ribosomal RNA genes ($N = 21$). For cyclopoid copepods, a total of 42 individuals were analyzed for the mitochondrial cytochrome c oxidase subunit I (COI) gene and 14 individuals were analyzed for the nuclear internal transcribed spacer 1 (ITS1) region. All 20 µL polymerase chain reactions contained 1x PCR buffer (Genscript), 1 mM MgCl$_2$, 1.0 µM of each primer, 0.08 µM dNTPs, 0.4 units of Taq DNA polymerase (Genscript), and 100 ng gDNA. Following amplification, PCR products were purified using the Solid Phase Reversible Immobilization (SPRI) method (DeAngelis et al. 1995). Sequencing reactions were performed in both directions using forward and reverse primers (Table 3.2) and BigDye Terminator 3.1 chemistry on an ABI 3130XL automated sequencer (Applied Biosystems, Foster City, CA).

**Phylogenetic and coalescent analyses**

DNA sequences for each marker were aligned, trimmed, and quality controlled using CodonCode Aligner v. 2.0.6 (CodonCode Corporation, Dedham, MA). For calanoids, the 18S and 28S alignments were concatenated following a non-significant incongruence length difference (ILD) test ($p = 1.0$; Farris et al. 1995), with 1000
homogeneity replicates and a heuristic search algorithm (with default settings) in PAUP* v. 4.0b.10 (Swofford 2002). Polymorphic sites and mutation biases were analyzed with DnaSp v. 5 (Librado and Rozas 2009). The best model of evolution for each dataset (calanoids: COI: TrN+I+G; ITS1: TVM+G; 28S+18S: TrN; cyclopoids: COI: TVM+G; ITS1: GTR+I+G), was determined using Modeltest v. 3.7 (Posada and Crandall 1998) under the Akaike Information Criterion (Posada and Buckley 2004). Neighbor-joining and Bayesian phylogenetic reconstructions were conducted for each marker. For the cyclopoid phylogenies, I also included the top BLAST matches from GenBank using the default blastn search algorithm to search for other closely related organisms. Neighbor-joining (NJ) phylogenetic reconstructions were performed in MEGA v. 4 (Tamura et al. 2007) using the TrN substitution model and 1000 bootstrap replicates. Bayesian (BI) reconstructions were conducted with MrBayes v. 3.1 (Ronquist and Huelsenbeck 2003) using the best fit substitution models as determined by Modeltest, and consisted of 4 replicate runs with 4 chains of 10,000,000 generations (1,000,000 for cyclopoids), discarding the first 25% as burn-in. The calanoid copepod *Leptodiaptomus siciloides*, collected from Lake Erie, Ontario, Canada, was used to root the calanoid phylogenies, while the harpacticoid copepod *Tigriopus japonicus* (accession nos. JN129181, EU057573) was used to root the cyclopoid trees.

I explored gene flow between the diaptomid populations of the three Malili lakes (Matano, Mahalona, and Towuti) using coalescent simulations of the full COI alignment and longest non-recombining stretch of DNA from the ITS1 alignment in the program IMa2 (Hey 2010). Pilot runs of 100,000 generations were conducted with priors for population sizes, divergence times, and migration rates chosen as per the user guidelines,
as well as priors 5x, 10x, and 15x larger. Small parameter autocorrelations, trendplots, and large ESS estimates indicated good chain mixing. In all pilot runs, the posterior probability distributions for some of the parameters varied and did not flatten to zero, indicating dependence on priors and low reliability of parameter estimates. However, the shapes of the probability distributions were similar between runs, indicating consistent general patterns in the data (i.e., non-zero migration rates). I conducted five final runs of 1,000,000 generations using the 5x priors. Log-likelihood ratio tests were conducted to infer statistical significance of migration between lakes Mahalona and Towuti. I used the following command lines for my analyses:

**M mode:**
```
ima2.exe -i datafull.txt -b 300000 -d 10 -hfg -hn 80 -ha 0.9 -hb 0.8 -l 100000 -m 4.5 -o datafulloutput.txt -p 2567 -r 25 -t 22 -q 54.5 -z 100
```

**L mode:**
```
ima2.exe -c 12 -i datafull.txt -b 300000 -l 100000 -m 4.5 -o lmodedatafull.txt -p 35 -r 0 -t 22 -q 54.5 -u 0.05 -w nestedmodels.txt -v test1datafull
```

**Demographic and population genetic analyses**

The number of haplotypes \((N_h)\), haplotype diversity \((h)\), nucleotide diversity \((\pi)\), Tajima’s \(D\) (Tajima 1989), and Fu’s \(F_S\) (Fu 1997) statistics were calculated for the major COI clades with DnaSp v.5 (Librado and Rozas 2009). Tajima’s \(D\) statistic (Tajima 1989) was used to test evolution under neutrality or demographic changes for each of the major COI clades. Significantly negative \(D\) values indicate purifying or directional selection or population expansion, while positive \(D\) values indicate recent bottlenecks or balancing selection. Population demographic changes were further investigated using Fu’s \(F_S\) (Fu 1997) and pairwise mismatch distributions computed with \(10^4\) permutations in Arlequin v. 3.5 (Rogers and Harpending 1992; Excoffier and Lischer 2010). Statistically
significant negative $F_S$ values and unimodal mismatch distributions indicate an excess of recent mutations and population expansion events, whereas positive $F_S$ values indicate lack of alleles or overdominant selection and multimodal distributions indicate demographic equilibrium. The relationships among the COI haplotypes were further examined by constructing a statistical parsimony haplotype network at the 95% connection limit in TCS v. 1.21 (Clement et al. 2000).

**Results: calanoid copepods**

*Morphological identification and genome size analysis*

Three morphological species of diaptomids were found across the island, each belonging to a different genus and inhabiting a geographically isolated lake system (Lake Tondano: *Phyllodiaptomus sulawesensis* Alekseev et Vaillant sp. nov.; Lake Poso: *Neodiaptomus lymphatus* Brehm, 1933; Lake Matano: *Eodiaptomus wolterecki* var. *matanensis* Brehm, 1933; Lakes Mahalona and Towuti: *Eodiaptomus wolterecki* Brehm, 1933; Table 3.1). Among the Malili Lakes, there were no major morphological differences between individuals. Genome sizes were constant within species, eliminating polyteny or polyploidy as a source of cryptic speciation (Table 3.1; Wyngaard and Rasch 2000).

*Sequence polymorphism and phylogenetic analyses*

The COI alignment consisted of 511 bp and 418 sequences and contained 179 synonymous and 4 nonsynonymous mutations comprising 150 unique haplotypes (Table 3.1). The mitochondrial phylogeny was used to select 49 representative individuals for sequencing the nuclear genes. The ITS1 alignment consisted of 708 bp and 49 sequences
and contained 154 variable sites and 10 unique haplotypes, while the 980 bp 18S+28S alignment consisted of 21 sequences with 36 variable sites and 4 unique haplotypes.

Phylogenetic tree topologies were consistent between neighbor-joining and Bayesian inference methods. The mitochondrial COI phylogeny reveals four well supported, highly divergent clades corresponding to populations in lakes Tondano (TD1 and 2), Poso (PO), Matano/Mahalona (MA/MH), and Towuti (TO) (Fig. 3.1b). Within Lake Tondano, there are two highly divergent lineages (15.4% sequence divergence; Table 3.4). Pons et al. (2010) conducted a detailed estimation of nucleotide substitution rates across the full mitochondrial genomes of 15 Coleopteran taxa and found rates for COI to be $0.08606 \pm 0.00338$ substitutions/site/million years/lineage or 17.8%–16.5% sequence divergence/my. However, substitution rates for COI are variable and thus lineage ages must be interpreted with caution. Based on this clock, the TD1 and TD2 lineages diverged approximately 860,000–930,000 years ago. A more traditional clock of 1.4% sequence divergence/my estimated from Alpheus shrimp around the Isthmus of Panama (Knowlton and Weigt 1998) suggest these lineages diverged 11 my ago. The population of Lake Poso constitutes a single monophyletic group. The Malili lakes contain two divergent clades, MA/MH and TO. Individuals belonging to the MA/MH clade are found in lakes Matano and Mahalona, with 4 haplotypes shared between both lakes. Lake Towuti harbours its own divergent lineage (12.7% sequence divergence from MA/MH; Table 3.4) corresponding to about 710,000–790,000 years of molecular divergence with a fast mtDNA clock, and 9 my ago with a slow clock. Only 3 individuals grouping with MA/MH were recovered at site TO1, the mouth of River Tominanga.
There was no evidence for intralacustrine population structure as individuals from different sampling locations were evenly distributed within each clade.

The nuclear phylogenies corroborate the divergences between genera, albeit with fewer informative sites because of highly conserved ribosomal sequences (Fig. 3.1c). However, there were two major discordances between the mitochondrial and nuclear trees. First, the two divergent mitochondrial clades in Lake Tondano (TD1 and 2) collapse to a single nuclear genotype. Second, both mitochondrial clades from the Malili lakes (MA/MH and TO) form a single nuclear clade with identical 18S and 28S sequences, and very small variation in ITS1 genotypes (0–0.6%). For coalescent simulations in IMa2, in all pilot runs and 5 final independent runs, migration parameters were consistently greater than zero for bidirectional migration between Matano and Mahalona and unidirectional migration from Mahalona into Towuti. A log-likelihood ratio test of the longer runs confirmed the rejection of a model with zero gene flow from Mahalona to Towuti (p = 0.016), indicating unidirectional gene flow from Lake Mahalona into Lake Towuti.

**Demographic and population genetic analyses**

Tajima’s $D$ values were non-significant within each of the major COI clades, indicating that the mitochondria is under purifying selection but otherwise evolving neutrally (Table 3.3). The population of Lake Tondano is genetically impoverished with only two major haplotypes (haplotype diversity; $h = 0.530$), indicative of a very recent and severe bottleneck or selective sweep. Populations of Poso and the Malili lakes had very high haplotype diversities ($h = 0.91–1.0$), suggesting high genetic diversity and large effective population sizes. Most notably, Lake Poso has a haplotype diversity of 1.0,
where every individual possesses a unique haplotype. Populations from Poso, Matano/Mahalona, and Towuti have significantly negative values of $F_S$ indicating an excess of rare haplotypes, likely the result of a recent population expansion and very large effective population sizes. This is corroborated by the normal mismatch distributions and complex web-like structures of the maximum parsimony networks of each clade (Fig. 3.2 and 3.4). In the MA/MH and TO clades there are several frequent haplotypes with many radiating branches, suggesting recent population expansion. The distribution of individuals from different sampling sites did not suggest any intralacustrine population structure, indicating that intralacustrine populations are homogenous and sampling was representative.

Results: cyclopoid copepods

Morphological identification

Cyclopoid copepods were present in lakes Poso, Matano, Mahalona, and Towuti. The species of Lake Poso was identified as *Thermocyclops crassus* Fischer, 1853. In the Malili lakes, I found two different genera restricted to specific lakes. In Lake Matano, the endemic *Tropocyclops matanoensis* Defaye, 2007 was identified. Lakes Mahalona and Towuti harbour species of the genus *Mesocyclops* Sars 1914, though only the population of Mahalona was identified to species level as *Mesocyclops aequatorialis similis* Van de Velde, 1984.

Phylogenetic analyses

The mitochondrial COI alignment consisted of 643 bp and 42 sequences and contained 146 synonymous and 79 nonsynonymous mutations and 22 unique haplotypes. The ITS1 alignment consisted of 722 bp and 14 sequences and contained 208 variable
sites and 8 unique haplotypes. The deep level phylogeny between genera was poorly resolved and incongruent between NJ and BI reconstructions. This is likely due to high levels of sequence divergence and subsequent loss of information at these loci. The mitochondrial COI phylogeny shows well-supported monophyletic clades for each of the lake populations (Fig. 3.3a). All of the species from Sulawesi group together and are extremely divergent from other cyclopoid species retrieved from GenBank. The species of Mahalona and Towuti are most similar, with 20% sequence divergence (approx. 1.2 my) corroborating their congeneric identification. The nuclear ITS1 phylogeny also reveals separate clades for each of the lake populations (Fig. 3.3b). However, there is strong support for reciprocal monophyly of the Mahalona and Towuti populations with 12% sequence divergence between them. They group closest to other *Mesocyclops* species, corroborating the morphological identification. However, *Thermocyclops crassus* from Lake Poso also groups here, albeit with high dissimilarity. The Lake Matano species is highly divergent from all other cyclopoid species, grouping closest to the parasitic copepod *Mytilicola* with low node support, confirming its long-term isolation and endemism.

**Discussion**

*Endemism, colonization, and local adaptation*

The morphological and phylogenetic analyses support the long-term isolation and endemism of Sulawesi’s fauna. All of the copepod species identified here appear to be endemic to Sulawesi, and the diaptomids appear to be of Asian origin, as nearly all of their congeners are inhabitants of Southeast Asia (Reddy 1994). Each of the cyclopoid species I found are highly divergent from other known cyclopoid species. The
colonization of each isolated lake system (Tondano, Poso, Malili) by different genera supports colonization order and allopatric speciation as the dominant force governing diversity over large geographic scales (>100 km).

Colonization order and natural selection is implicated as an important factor in determining species distributions. First, each lake harbours only a single calanoid and cyclopoid species. This is not likely the consequence of extremely rare colonization events, as zooplankton have high dispersal abilities. More likely, strong natural selection and local adaptation prevents the establishment of ecologically similar species through competitive exclusion (see Louette and De Meester 2007). Thus, colonization order of lake systems is an important component of planktonic speciation. Second, copepod populations are highly divergent between adjacent lakes despite the high potential for dispersal. The *Mesocyclops* populations of lakes Mahalona and Towuti are reciprocally monophyletic and probably derived from a common ancestor, yet are highly divergent (COI: 20% sequence divergence; ITS1: 12% sequence divergence; Fig. 3.3).

Furthermore, Lake Matano harbours another endemic species from a different genus which remains restricted to the lake. In diaptomids, highly divergent mitochondrial clades were found between the populations of lakes Matano and Mahalona and Lake Towuti. These distributions reveal strong reproductive/dispersal barriers between each of the Malili lakes and indicate that strong local adaptation and natural selection facilitate divergence over very short geographic ranges. The divergence time estimates of 710,000–790,000 years between the diaptomid MA/MH and TO clades corresponds roughly to the approximate age of Lake Towuti, which is thought to be at least 600,000 years old, suggesting the clades have been diverging since the lake’s formation (Vaillant
et al. 2011). About one third of the Malili lakes’ diatom species are endemic to single lakes and the physicochemical environments vary among the lakes (Bramburger et al. 2008; see Appendix I). Thus, there is certainly the potential for local adaptation of copepods to different physical conditions and food sources.

The population genetic analyses reveal some evidence for demographic changes in the diaptomids. The population of Lake Tondano shows evidence of an extreme population bottleneck or selective sweep, as indicated by a severe impoverishment of genetic diversity and a lack of rare alleles. This may be the result of drastic changes in environmental conditions within the last 33,000 years including local volcanic eruptions and rapid lake level fluctuations (Dam et al. 2001). By contrast, Lake Poso and the Malili lakes populations have high haplotype diversities, low \( F_S \) values, unimodal mismatch distributions, and complex haplotype networks with star-like branching patterns (Table 3.1; 3.3; Fig. 3.2). This indicates very high genetic diversity, large effective population size, and some population expansion – a testament to the prolonged existence of the lakes. In both Lake Tondano and the Malili lakes, demographic factors have influenced the evolutionary trajectory of the diaptomid populations, though the relative contributions of genetic drift and selection are uncertain.

The Malili lakes harbour several intralacustrine adaptive radiations of fish, shrimp, gastropods, and crabs (reviewed in Vaillant et al. 2011; von Rintelen 2011). However, the copepods have not undergone intralacustrine adaptive radiations like these other taxa. It has been proposed that this lack of radiation in zooplankton is primarily because the homogenous pelagic environment prevents niche diversification (Dumont
My results are consistent with other ancient lakes, which tend to harbour few species of copepods (Martens and Schon 1999).

**Hybridization in diaptomid populations**

In many copepods, large mitochondrial sequence divergences between morphologically identical populations are common, typically ranging from 0–3% within populations and up to 15–25% between isolated populations or congeners over regional geographic ranges (Adamowicz et al. 2007; Makino and Tanabe 2009; Thum and Harrison 2009; Makino et al. 2010; Marrone et al. 2010). The phylogenetic analyses of diaptomids reveal discordance between the mitochondrial and nuclear trees. Both Lake Tondano and the Malili lakes harbour a pair of cryptic, divergent mitochondrial clades (TD1 and TD2, 15.4% sequence divergence; MA/MH and TO, 12.7% sequence divergence; Fig. 3.1b). However, each pair of clades shares nuclear alleles, suggesting gene-flow between them. There are several potential explanations for this cytonuclear discordance. First, limited dispersal of females could create strong geographic structuring in the mitochondria, however both sexes were found in the rivers connecting the Malili lakes. Second, a recent selective sweep on mitochondria or other maternally inherited elements (e.g., endosymbionts; Hurst and Jiggins 2005) can rapidly replace the mitochondrial lineage of a population. However, this is unlikely in the Malili lakes populations because there is no evidence of a genetic bottleneck, and unlikely in the Tondano population because the sequence divergence between mitochondrial lineages is too large to have recently arisen from the same population. Third, the similarity of nuclear alleles is due to an exceptionally slow evolutionary rate for these nuclear loci (i.e. incomplete lineage sorting), however this unlikely as other copepods of similar
mitochondrial divergence usually show fixed polymorphisms in nuclear ribosomal genes (see Rocha-Olivares et al. 2001; Thum and Harrison 2009; Machida and Tsuda 2010; Makino et al. 2010). For example, between cryptic *Neocalanus* species with 15.4–16.2% divergence in COI, there are 12 polymorphic sites in the nuclear 28S and ITS1 regions (0.8–1.1% and 0.4–0.7% divergence, respectively) which are all diagnostic of species (Machida and Tsuda 2010). In the Malili Lakes, there are 7 polymorphic sites in 18S, 28S, and ITS1, none of which distinguish between the MA/MH and TO mitochondrial clades. The most likely scenario is that individuals belonging to these clades are hybridizing and the ribosomal genes are homogenizing through concerted evolution due to high rates of unequal crossing over and gene conversion (Hillis and Dixon 1991). This high rate of homogenization combined with low sampling explains the lack of detection of hybrid individuals. Many studies of copepods show that crosses between congeners, some with as much as 22% mitochondrial sequence divergence, can result in the production of hybrids, indicating that prezygotic isolation evolves slowly and large mitochondrial sequence divergences are not always indicative of reproductive isolation (Burton 1990; Byrom et al. 1993; Chen et al. 1997; Harrison and Edmands 2006; Thum 2007; Thum and Harrison 2009).

The phylogenetic results suggest that the diaptomid population of Lake Tondano is of hybrid origin from two recent colonists or two previously isolated parental lineages. In this population, it appears that neither parental mitochondrial lineage has become fixed, but nuclear ribosomal genes have homogenized through gene flow and recombination. In the diaptomid populations of the Malili lakes, the MA/MH mitochondrial clade is restricted to the northern lakes Matano and Mahalona while the
TO clade is restricted to southern Lake Towuti despite the strong north-south water flow and high dispersal ability of copepods. By contrast, nuclear alleles are shared between the clades. I suggest that unidirectional hybridization is responsible for restricting the geographic distribution of mitochondrial clades, but allows gene flow between the nuclear genomes. This is supported by the log-likelihood ratio test of the coalescent simulations which detected unidirectional gene flow between the MA/MH and TO clades. Unidirectional hybridization can result from i) cytonuclear incompatibilities due to dysfunctional interactions between uniparentally inherited mitochondria and the nuclear genes from divergent parental lineages (Turelli and Moyle 2007), ii) hybrid breakdown of the heterogametic sex (e.g., XY, WZ) (Haldane’s rule), or iii) other factors such as divergence in mate preference or mating behaviours that result in asymmetrical hybridization (Wirtz 1999). Cytonuclear incompatibilities seem the most likely explanation for this pattern in the Malili populations, as this mechanism has been well demonstrated as the cause of asymmetrical hybrid breakdown in the harpactocoid copepod *Tigriopus californicus* (Ellison and Burton 2008). Haldane’s rule is unlikely to create the observed pattern because sex chromosomes are absent in most copepods (Dussart and Defaye 2001). Copepods exhibit a wide array of sexual behaviours which could also be a cause of reproductive barriers (Dussart and Defaye 2001). It is not difficult to imagine small variations in morphology, mate recognition, capture, spermatophore placement and attachment, and sperm competition between multiple matings as the cause of unidirectional hybridization. Thus, the most important forms of reproductive isolation between the diaptomid clades are habitat isolation between lakes, intrinsic postzygotic isolation from genetic incompatibilities, and perhaps behavioural or
mechanical prezygotic isolation from divergent reproductive characters. While hybridization has an important role in the evolution of diaptomids in Lake Tondano and the Malili lakes, the ultimate outcome of this process remains unclear.

**Conclusions**

My study demonstrates that various speciation processes operate at different spatial scales. At the island scale, colonization order and allopatric speciation are the rule. It has been demonstrated that large, locally adapted zooplankton populations can limit the establishment of subsequent colonists generating a strong priority effect (Louette and De Meester 2007). By contrast, at the intralacustrine scale, speciation processes are complex and may be influenced by strong natural selection, demographic changes, and hybridization with neighboring populations. Hybridization appears to be much more common than expected and the outcome depends on a balance between extrinsic (e.g., selection, environmental stochasticity) and intrinsic (e.g., co-adapted gene complexes, cytonuclear interactions, genomic rearrangements) factors (Liou and Price 1994; Dowling and Secor 1997; Mallet 2007; Pritchard et al. 2011). It has been suggested that low rates of gene flow between partially isolated, differentiated populations may accelerate speciation by providing an influx of novel genotypes and greater phenotypic diversity (Seehausen 2004). Because freshwater plankton are often partially isolated in neighboring waterbodies with intermittent dispersal between populations, hybridization may in fact be one of the most common and significant evolutionary forces acting in this group.
Figure 3.1: (a) A map of sampling sites across Sulawesi, Indonesia. The dashed arrows indicate the direction of water flow. The Bayesian phylogeny for the diaptomid COI gene (b) shows two divergent mitochondrial lineages inhabiting Lake Tondano and two in the Malili lakes. These lineages collapse in the Bayesian phylogeny for the diaptomid ITS1 region (c), indicating hybridization has homogenized the ribosomal genes. Numbers in parentheses indicate the number of sampled individuals within the clade or haplotype. Node supports are posterior probabilities followed by bootstrap values.
Figure 3.2: Maximum parsimony haplotype network for the major COI clades. Colored circles represent haplotypes, with size corresponding to haplotype frequency. Single lines correspond to single mutation steps (i.e., one base-pair change) and small open circles represent extinct or unsampled haplotypes.
Figure 3.3: The Bayesian phylogeny for the cyclopoid COI gene (a) shows different mitochondrial lineages inhabiting Lake Poso and each of the Malili lakes. These lineages are supported in Bayesian phylogeny for the ITS1 region (b), indicating each species is distinct. Numbers in parentheses indicate the number of sampled individuals within the clade or haplotype. Node supports are posterior probabilities followed by bootstrap values.
**Figure 3.4:** Observed (open circles) and expected (closed circles) pairwise mismatch distributions COI sequences of diaptomid populations from Lake Poso and the Malili lakes. Unimodal distributions indicate population growth.
<table>
<thead>
<tr>
<th>Species name / Sampling Site</th>
<th>ID</th>
<th>Latitude</th>
<th>Longitude</th>
<th>N</th>
<th>N&lt;sub&gt;h&lt;/sub&gt;</th>
<th>h</th>
<th>π</th>
<th>Mean genome size (pg ± SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phyllodiaptomus sulawensis</strong></td>
<td>TD</td>
<td>1°13'02.57&quot;N</td>
<td>124°54'05.59&quot;E</td>
<td>59</td>
<td>4</td>
<td>0.530</td>
<td>0.068</td>
<td>0.81 ± 0.02</td>
</tr>
<tr>
<td>Tondano East</td>
<td>TD1</td>
<td>1°13'09.03&quot;N</td>
<td>124°53'39.65&quot;E</td>
<td>24</td>
<td>2</td>
<td>0.507</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>Tondano Center</td>
<td>TD2</td>
<td>1°13'09.11&quot;N</td>
<td>124°53'06.57&quot;E</td>
<td>20</td>
<td>4</td>
<td>0.574</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>Tondano West</td>
<td>TD3</td>
<td>1°13'09.11&quot;N</td>
<td>124°53'06.57&quot;E</td>
<td>15</td>
<td>2</td>
<td>0.533</td>
<td>0.072</td>
<td></td>
</tr>
<tr>
<td><strong>Neodiaptomus lymphatus</strong></td>
<td>PO</td>
<td>1°48'05.36&quot;S</td>
<td>120°37'00.38&quot;E</td>
<td>34</td>
<td>34</td>
<td>1.0</td>
<td>0.012</td>
<td>0.67 ± 0.03</td>
</tr>
<tr>
<td>Poso North</td>
<td>PO1</td>
<td>1°54'09.69&quot;S</td>
<td>120°37'01.91&quot;E</td>
<td>5</td>
<td>5</td>
<td>1.0</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>Poso Center</td>
<td>PO2</td>
<td>1°54'09.69&quot;S</td>
<td>120°37'01.91&quot;E</td>
<td>22</td>
<td>22</td>
<td>1.0</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>Poso South</td>
<td>PO3</td>
<td>2°00'07.51&quot;S</td>
<td>120°40'06.39&quot;E</td>
<td>7</td>
<td>7</td>
<td>1.0</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td><strong>Eodiaptomus wolterecki var. matanensis</strong></td>
<td>MA</td>
<td>2°27'00.79&quot;S</td>
<td>121°15'00.33&quot;E</td>
<td>99</td>
<td>37</td>
<td>0.947</td>
<td>0.006</td>
<td>0.75 ± 0.01</td>
</tr>
<tr>
<td>Matano Northwest</td>
<td>MA1</td>
<td>2°28'01.18&quot;S</td>
<td>121°18'00.93&quot;E</td>
<td>26</td>
<td>15</td>
<td>0.945</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Matano Center</td>
<td>MA2</td>
<td>2°30'02.66&quot;S</td>
<td>121°24'09.99&quot;E</td>
<td>42</td>
<td>19</td>
<td>0.906</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>Matano Southeast</td>
<td>MA3</td>
<td>2°32'06.07&quot;S</td>
<td>121°28'08.50&quot;E</td>
<td>15</td>
<td>9</td>
<td>0.867</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Petea outlet</td>
<td>MA4</td>
<td>2°33'16.30&quot;S</td>
<td>121°31'23.92&quot;E</td>
<td>4</td>
<td>3</td>
<td>1.0</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td><strong>Eodiaptomus wolterecki</strong></td>
<td>MH</td>
<td>2°34'24.61&quot;S</td>
<td>121°30'29.96&quot;E</td>
<td>96</td>
<td>50</td>
<td>0.936</td>
<td>0.011</td>
<td>0.74 ± 0.03</td>
</tr>
<tr>
<td>River Petea</td>
<td>MH1</td>
<td>2°34'33.90&quot;S</td>
<td>121°30'16.23&quot;E</td>
<td>22</td>
<td>10</td>
<td>0.814</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Petea mouth</td>
<td>MH2</td>
<td>2°35'16.60&quot;S</td>
<td>121°29'36.40&quot;E</td>
<td>25</td>
<td>14</td>
<td>0.893</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Mahalona East</td>
<td>MH3</td>
<td>2°35'43.74&quot;S</td>
<td>121°28'52.06&quot;E</td>
<td>15</td>
<td>9</td>
<td>0.905</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Mahalona Center</td>
<td>MH4</td>
<td>2°36'12.60&quot;S</td>
<td>121°31'32.51&quot;E</td>
<td>16</td>
<td>11</td>
<td>0.908</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Mahalona West</td>
<td>MH5</td>
<td>2°37'08.43&quot;S</td>
<td>121°33'01.92&quot;E</td>
<td>17</td>
<td>12</td>
<td>0.941</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td><strong>Eodiaptomus wolterecki</strong></td>
<td>TO</td>
<td>2°39'58.10&quot;S</td>
<td>121°34'08.95&quot;E</td>
<td>130</td>
<td>44</td>
<td>0.913</td>
<td>0.012</td>
<td>0.73 ± 0.02</td>
</tr>
<tr>
<td>Tominanga mouth</td>
<td>TO1</td>
<td>2°42'05.83&quot;S</td>
<td>121°35'02.41&quot;E</td>
<td>16</td>
<td>11</td>
<td>0.908</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Towuti North</td>
<td>TO2</td>
<td>2°48'08.43&quot;S</td>
<td>121°33'01.92&quot;E</td>
<td>17</td>
<td>12</td>
<td>0.941</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>Towuti East</td>
<td>TO3</td>
<td>2°49'08.95&quot;S</td>
<td>121°27'06.55&quot;E</td>
<td>84</td>
<td>27</td>
<td>0.902</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Towuti West</td>
<td>TO4</td>
<td>2°50'01.78&quot;S</td>
<td>121°27'06.55&quot;E</td>
<td>84</td>
<td>27</td>
<td>0.902</td>
<td>0.009</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.1:** Mitochondrial DNA genetic diversity indices and genome sizes of diaptomid populations from Sulawesi. N, number of individuals; N<sub>h</sub>, number of haplotypes; h, haplotype diversity; π, nucleotide diversity.
Table 3.2: Primer pairs and thermal cycling programs used for the COI, ITS1, 18S, and 28S amplification. Locality names are defined in Table 3.1.
Table 3.3: Results of diaptomid population demographic analyses for Tajima’s $D$, Fu’s $F_S$, and mismatch distributions. Significant values are highlighted in bold.

<table>
<thead>
<tr>
<th>Clade</th>
<th>Tajima’s $D$</th>
<th>Fu’s $F_S$</th>
<th>Mismatch distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D$</td>
<td>$p$</td>
<td>$F_S$</td>
</tr>
<tr>
<td>TD1</td>
<td>n/a</td>
<td>-</td>
<td>n/a</td>
</tr>
<tr>
<td>TD2</td>
<td>$-1.513$</td>
<td>$&gt; 0.10$</td>
<td>$-2.176$</td>
</tr>
<tr>
<td>PO</td>
<td>$-1.710$</td>
<td>$&gt; 0.05$</td>
<td>$-34.447$</td>
</tr>
<tr>
<td>MA/MH</td>
<td>$-1.551$</td>
<td>$&gt; 0.05$</td>
<td>$-86.316$</td>
</tr>
<tr>
<td>TO</td>
<td>$-0.397$</td>
<td>$&gt; 0.10$</td>
<td>$-20.257$</td>
</tr>
<tr>
<td></td>
<td>TD1</td>
<td>TD2</td>
<td>PO</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>TD1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TD2</td>
<td>0.154</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>PO</td>
<td>0.205</td>
<td>0.204</td>
<td>0.012</td>
</tr>
<tr>
<td>MA/MH</td>
<td>0.259</td>
<td>0.220</td>
<td>0.237</td>
</tr>
<tr>
<td>TO</td>
<td>0.268</td>
<td>0.233</td>
<td>0.222</td>
</tr>
</tbody>
</table>

**Table 3.4:** Mean corrected p-distances between major diaptomid COI clades. Italic values along the diagonal are average within-group distances.
References


CHAPTER IV
GENERAL DISCUSSION

Final summary

In this thesis, I present insights into speciation processes by investigating the species-flocks of the Malili Lakes of Sulawesi, Indonesia. I embarked on this endeavour by first reviewing the evolutionary research conducted in the Malili Lakes and highlighting trends in speciation patterns. Many of the taxa in the Malili lakes have undergone adaptive radiations with respect to trophic specialization, developing an array of morphological and behavioural traits to exploit specific resources. This fact, in addition to the extreme oligotrophy of the lakes, suggests that bottom up processes are responsible for high levels of speciation. That is, intraspecific competition for extremely limited resources has driven taxa to specialize for specific food sources. Furthermore, hybridization between closely related lineages appears to be an important evolutionary process. While its role remains somewhat uncertain and probably varies between taxa, it is clear that hybridization is common and likely drives diversification by increasing phenotypic diversity within populations.

In the third chapter, I explored the evolutionary histories of copepod populations from Sulawesi. I found that colonization order is likely responsible for the distribution of zooplankton in freshwater ecosystems. Priority effects (e.g., competitive exclusion) then maintain these distributions and prevent further colonization of ecologically similar species (De Meester 2002). Consistent with other Malili lakes species-flocks, I also found that hybridization may be a common phenomenon in planktonic taxa and plays an important role in their speciation. Furthermore, divergence was only found in the
mitochondrial genome, demonstrating that some genomic elements can diverge while others remain homogenous between differentiated hybridizing lineages. This is consistent with the idea of “porous” genomes which suggests that the genomes of hybridizing species can form a mosaic of divergent and homogenous regions. Disruptive selection can drive divergence of particular genes, their surrounding regions (divergence hitchiking), and other interacting elements (e.g., mitochondria), while selectively neutral loci are free to recombine and homogenize (Gavrilets and Vose 2005; Via and West 2008). This can result in species that remain distinct and separate despite consistent hybridization between them.

**Closing thoughts**

One of the overarching questions in speciation research is why adaptive radiations occur so often in ancient lakes. I argue that high rates of speciation in these systems are driven by a unique interaction between environment and genetics that is absent in other types of ecosystems. First, the environment of ancient lakes is extremely conducive to speciation. These long-lived lakes have undergone substantial environmental fluctuations over much longer timescales than other young lakes (see Cristescu et al. 2010). Thus, ancient lakes represent persistent, isolated habitats that are in constant change. Marine and continental species often respond to shifting environmental conditions by shifting their geographic distributions (Davis and Shaw 2001; Beaugrand et al. 2002). In ancient lakes however, species are well adapted to the local environment have little opportunity to disperse to ecologically similar habitats (i.e., refugia). Environmental fluctuations often resulted in alternating periods of sympatry and allopatry, population bottlenecks, and varying selective pressures as habitats changed (e.g., Rüber et al. 1999; Cristescu et
Over geological timescales, such environmental variation would either drive species to extinction, or select for species with a high capacity to adapt to changing and unpredictable conditions. I suspect that many ancient lake species have evolved mechanisms to enhance their capacity for local adaptation (e.g., higher mutation rates, phenotypic plasticity, strong habitat preference), though this remains to be tested. The same adaptations would also be conducive to speciation.

The second factor important for driving speciation is a genetic predisposition to speciate. This is in part determined by genetic architecture, the physical arrangement of elements in the genome (Bell and Travis 2005). Thus, different taxa carry with them an intrinsic tendency toward speciation, evidenced by the occurrence of multiple adaptive radiations in the same lineage (Seehausen 2006). As previously discussed, species in ancient lakes are subject to repeated periods of sympatry and allopatry, founder effects, and varying selection regimes. Over time, this process produces many differentiated lineages which come into secondary contact. Gene flow between these lineages has the potential to increase phenotypic diversity through transgressive segregation, the appearance of extreme or novel phenotypes (Seehausen 2004; Bell and Travis 2005; Mallet 2007). This may increase the probability of speciation through two main mechanisms: 1) the production of hybrids with higher phenotypic diversity, facilitating colonization of new niches, and 2) introgressive hybridization, resulting in genomic changes in hybrids or parental populations which make them more inclined to speciate (i.e., through new linkages between mate preference, habitat preference, sexual, and ecological traits). The combination of long-term environmental heterogeneity and
intermittent hybridization between lineages in ancient lakes creates a unique situation where speciation is common.

There are also several interesting avenues for future research in this area. One potential is to confirm the prevalence of hybridization among freshwater zooplankton populations with a meta-analysis of multilocus studies over a broader taxonomic and geographic range. Widespread hybridization would demonstrate the pervasiveness of overlapping species boundaries in taxa which are distributed as metapopulations. Furthermore, this would highlight the importance of hybridization as a creative evolutionary force.

Another research area worthy of exploration is the relationship between gene flow and phenotypic diversity. While many studies now show that introgressive hybridization increases phenotypic diversity in hybrids as well as parental populations, it is difficult to quantify this relationship. It may be possible to compare speciation rates of species-flocks that show evidence for hybridization to those that do not hybridize. Coalescent simulations or population genetics statistics (e.g., $F_{ST}$) could be used as a proxy for actual gene flow between populations (though this is locus dependent). Is there a negative correlation between gene flow and speciation rate (i.e., does gene flow always inhibit divergence? Fig. 4.1a), or is there a maximum, where intermediate gene flow accompanies the highest speciation rates (i.e., the production of new, stable evolutionary lineages; Fig. 4.1b)? Related questions include how much gene flow maximizes phenotypic diversity in hybrids or parental populations, and how much initial genetic divergence between populations maximizes phenotypic diversity when they do hybridize. These questions may prove challenging, as the answers certainly depend on a myriad of
variables including mutation rates, genomic architecture (i.e., the configuration of genes in the genome), cytonuclear interactions (e.g., mitochondria, endosymbionts), epigenetic effects, and variability of the environments of different populations. Nevertheless, general trends across taxa and loci may appear.

In conclusion, ancient lakes represent ideal systems to study and explore speciation. Processes like long-distance dispersal, allopatric speciation, genetic drift, ecological and sexual selection, and hybridization all act at various geographic and temporal scales. Environmental factors also profoundly influence these processes, from resource limitation to climatic fluctuation. Future research in these remarkable habitats will surely reveal much about the origins of biodiversity on earth.
**Figure 4.1:** Possible relationships between speciation rates and gene flow between populations. Populations experiencing complete admixture (i.e., Hardy-Weinberg equilibrium) show a baseline phenotypic diversity and no speciation. As gene-flow between two populations decreases, total phenotypic diversity and the probability of speciation increases until there is complete isolation between populations and they diverge at a neutral rate (as a function of mutation rates, genetic drift, and natural selection in their local environments). Gene flow may always act as a barrier to divergence by homogenizing genomes (a), but I hypothesize that there may be an intermediate level of gene flow where phenotypic diversity and probability of speciation is maximized through the exchange of novel alleles and transgressive segregation (b).
References


## APPENDIX I

### PHYSICAL AND CHEMICAL CHARACTERISTICS OF THE MALILI LAKES

<table>
<thead>
<tr>
<th>Lake</th>
<th>Age (approx.)</th>
<th>Area (km²)</th>
<th>Max. Depth (m)</th>
<th>Metal concentrations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ba (µg/L)</td>
</tr>
<tr>
<td>Matano</td>
<td>2-4 myr</td>
<td>164</td>
<td>590</td>
<td>surface</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50 m</td>
</tr>
<tr>
<td>Mahalona</td>
<td>&lt;500,000 yr</td>
<td>25</td>
<td>62</td>
<td>surface</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30 m</td>
</tr>
<tr>
<td>Towuti</td>
<td>~600,000 yr</td>
<td>560</td>
<td>200</td>
<td>surface</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50 m</td>
</tr>
</tbody>
</table>

Selected physical and chemical characteristics of the Malili lakes. Measurements were taken in June, 2009.
<table>
<thead>
<tr>
<th>Name</th>
<th>James Joseph Vaillant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birthdate</td>
<td>1987</td>
</tr>
<tr>
<td>Birthplace</td>
<td>Sarnia, Ontario, Canada</td>
</tr>
<tr>
<td>Education</td>
<td>St. Christopher Secondary School 2001–2005</td>
</tr>
<tr>
<td></td>
<td>B.Sc. – University of Windsor 2005–2010</td>
</tr>
<tr>
<td></td>
<td>M.Sc. – University of Windsor 2010–2012</td>
</tr>
</tbody>
</table>