ABSTRACT

Evolutionary Forces Driving Population Differentiation in Lake Malawi Rock-Dwelling Cichlids (Pisces: Cichlidae)

Martin Husemann, Ph.D.

Mentor: Patrick D. Danley, Ph.D.

The East African cichlids with more than 2000 species represent the most diverse vertebrate radiation known. Lake Malawi harbors the most species rich flock with more than 700 endemic cichlids. In this work I use a population based approach to study the forces driving the divergence of populations and the factors contributing to the maintenance of species diversity. I examine the effect of genetic drift on population divergence through time and space using an analysis of effective population sizes. This study indicates that populations of the widespread *Maylandia zebra* are over 550 individuals in size and are at an equilibrium state. The microendemic *Maylandia benetos* has a relatively small population size (~500 individuals) and evidence for drift is found. I also examine the phenotypic divergence in ecological and sexual characters in *M. zebra*. This study reveals that ecological selection plays an important role in the divergence of body shape and length in *M. zebra*. Divergence in male color pattern shows weaker evidence for selection. Furthermore, the divergence estimates for body shape and melanophore count are slightly correlated. This correlation between an ecological and a

sexual trait may indicate that both phenotypes might be under correlational selection. Correlational selection on ecological and sexual traits is further supported by the correlation of body shape and male coloration in replicated sympatric species pairs of *Maylandia*. A common garden experiment using the sympatric *M. zebra* and *M. benetos,* representing a sympatric species pair with divergent coloration, shows that body shape had a strong genetic component but also exhibited phenotypic plasticity. Hybrid crosses of the two species reveal a complicated mode of inheritance for body shape differences and demonstrated high degrees of transgressive segregation. Overall, selection appears to be the driving force of phenotypic evolution in rock-dwelling cichlids. However, small population sizes of microendemics expose them to the effects of drift. Evidence for correlational selection suggests that specific combinations of reproductive and ecological traits might be favored. Finally, phenotypic plasticity and transgressive segregation are two mechanisms generating new phenotypic diversity contributing to the diversification of cichlids.

Evolutionary Forces Driving Population Differentiation in Lake Malawi Rock-Dwelling Cichlids (Pisces: Cichlidae)

by

Martin Husemann, B.Sc, M.Sc.

A Dissertation

Approved by the Department of Biology

_____

Robert D. Doyle, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Doctor of Philosophy

Approved by the Dissertation Committee

_____

Patrick D. Danley, Ph.D., Chairperson

_____

Daniel J. Peppe, Ph.D.

_____

Darrell S. Vodopich, Ph.D.

_____

Ryan S. King, Ph.D.

_____

Stephen J. Trumble, Ph.D.

_____

Hans A. Hofmann, Ph.D.

Accepted by the Graduate School

December 2013

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I thank my advisor and my committee: Dr. Patrick D. Danley, Dr. Daniel Peppe, Dr. Ryan S. King, Dr. Darrell Vodopich, Dr. Stephen Trumble and Dr. Hans Hofmann for their help and encouragement. I would like to thank the members of the Danley Lab for fruitful discussions and ideas. Thanks to Richard Zatha and the Malawi field crew for help with sampling the fish used in this study. I am grateful to the officers of Lake Malawi National Park, members of Malawi Fisheries Department, and faculty at the University of Malawi for assistance in conducting the necessary field work.

I am especially grateful to Baoqing Ding and Dr. Michael Tobler for the help and advice with all kinds of analyses and their input to my studies. Thanks to the undergraduate students, Cagney McCauley, Rachel Nguyen, Suk Namkung, Cecilia Munoz-Ducoing, I had the opportunity to work with and who contributed a lot of data to this dissertation.

Thanks to my friends, Mo, Eun-Jeong, Elli, Michael, and Baoqing, who made graduate school fun. Most of all I am very grateful to my family who always had my back and constantly supported me. I could not have done it without you. Miriam, I am so lucky to have you. Thanks for being at my side during all of this supporting me with your unconditional love.

Funding for my dissertation was provided by the American Cichlid Association, Glasscock Fund for Excellence in Environmental Science, the Jack G. and Norma Jean

Folmar Research Fund, and a Bob Gardner Memorial grant. The Biology Department always supported me when in need.

DEDICATION

Für meine Eltern

CHAPTER ONE

Introduction

A main goal of biology is to understand how species evolve. A first step in understanding speciation is to estimate the factors driving the divergence of populations. Generally three main evolutionary forces need to be considered on ecological time scales: drift, migration and selection. The relative role each of these forces plays depends on species specific intrinsic characteristics as well as the landscape's composition and history (e.g. Aguilée et al. 2011, Husemann et al. 2012a).

Drift

Genetic drift is considered the null model of population divergence. Drift is the result of random processes in finite populations and will lead to the fixation of alleles if not counteracted by migration or mutation. Hence, the strength of drift strongly depends on the effective size of a population, gene flow from other populations and the time scale over which a population is observed. However, drift is thought to be of limited importance over ecological timescales if effective population sizes are larger than 500 individuals (Hartl & Clark 2007). However, in small and isolated populations drift can lead to significant fluctuations of allele frequencies and the fixation of alleles even over short time scales. Therefore, in the absence of gene flow drift can lead to rapid divergence if populations are small (Ellstrand & Elam 1993).

Migration

Migration and gene flow, however, counterbalance the effects of drift by introducing new alleles into a population. The rate of gene flow depends on the effective number of migrants a population receives and the mating success of immigrants. Migration rates generally depend strongly on the landscape and the habitat matrix between patches of suitable habitat (Ricketts 2001, Manel & Holderegger 2013). In widespread taxa the connectivity of subpopulations often leads to a meta-population structure with source and sink dynamics (Hanski 1999, Waits et al. 2008). In such meta-population networks some populations serve as source populations which have the highest diversity and usually are located in the largest and highest quality stretches of habitat. The sink populations are smaller and receive migrants from the source. Populations well connected in such networks are receiving new alleles from migrants at the same rate as they loose alleles to drift and therefore are in migration-drift equilibrium.

Selection

The balancing effect of gene flow can be overcome if the selective environments of populations differ (Danley et al. 2000, Nosil 2008). In such cases divergence can occur even in the presence of gene flow. Hence, selection can be considered the most powerful of the evolutionary forces (Darwin 1859). Selection is a directional force and drives populations to a local optimum of a trait. Different kinds of selection can be distinguished depending on which kind of character it acts on and in which direction the evolution of a trait is driven. Directional selection drives a trait to a specific optimum, stabilizing selection keeps the trait at an optimum and disruptive selection can lead to multiple character states at different optima. Further, natural and sexual selection can be

distinguished depending on the type of trait on which selection acts. Sexual selection acts on traits involved in mating success, whereas natural selection drives the divergence of ecological characters.

Modes of Speciation

If selection and drift can overcome the homogenizing effects of migration speciation might occur. While there is a large suite of settings which can generate such conditions often geographic barriers and divergent ecological conditions play dominant roles in the divergence of populations and species. Historically, modes of speciation were grouped based on the strength of geographic barriers. Allopatric speciation occurs when populations are geographically isolated. Parapatric speciation occurs when populations are adjacent to each other but occupy different habitats. This mode of speciation is usually associated with environmental gradients. Sympatric speciation occurs in the absence of a geographic barrier. As sympatric speciation is difficult to prove, the term ecological speciation has been coined for situations in which no geographic barrier is obvious and species diverged as a result of ecological selection (Schluter 2001). In many cases, however, the distinction between different modes is not easy to make; for example, populations might have initially diverged in allopatry but speciation was only completed when incipient species came into secondary contact and evolutionary mechanisms such as reinforcement and character displacement led to reproductively isolated and ecologically distinct species. Again, understanding the forces that have led to the initial population divergence might help to understand what has caused speciation.

Species Radiations as Model Systems

If the geographical setting and the environmental conditions are suitable, species groups can sometimes undergo a rapid diversification process which results in a species radiation (Aguilée et al. 2011, Danley et al. 2012a, Nevado et al. 2013). Often such radiations are found in mosaic like habitats such as archipelagos, mountain ranges, or lakes with scattered distributions of habitat types. Often rapid speciation is then driven by ecological opportunity where species of the same group adapt to a variety of ecological niches (Gavrilets & Losos 2009). In other cases radiations might be non-adaptive, largely driven by allopatric processes and ecological divergence occurs subsequently to speciation (Rundell & Price 2009). In either case species radiations present researchers with large numbers of species in different ecological settings and therefore provide natural laboratories for speciation research.

Common examples for such radiations are the *Anolis* lizards of the Caribbean (Losos et al. 2006), the fruit flies and crickets on Hawaii (Zimmerman 1970, Hoy et al. 1988, Shaw 2002), the finches on Galapagos (Grant 1981, Grant & Grant 2006), *Melanoplus* grasshoppers on the sky islands of North America (Knowles 2001, Carstens & Knowles 2007), *Buthus* scorpions in the North African Atlas Mountains (Habel et al. 2012, Husemann et al. 2012b) Amphipods in Lake Baikal (Sherbakov 1999, McDonald 2005), and East African cichlids (Kornfield & Smith 2000, Kocher 2004, Genner & Turner 2005, Seehausen 2006, Salzburger 2009, Sturmbauer et al. 2011, Danley et al. 2012a).

The Cichlid System

The East African cichlids, with more than 2000 species, represent the most diverse vertebrate radiation known (Kocher 2004, Genner & Turner 2005, Sturmbauer et al. 2011, Danley et al. 2012a). Most species are found in the three East African Great Lakes: Lakes Victoria, Tanganyika and Malawi. Phylogenetic studies have revealed that the species flock originated in Lake Tanganyika from where it colonized the other lakes (Salzburger et al. 2005). The radiations within the lakes were strongly influenced by the geological and climatic history of the region and diversification occurred in response to large lake level fluctuations (Sturmbauer et al. 2001, Genner et al. 2010, Danley et al. 2012a, Nevado et al. 2013).

In addition to these extrinsic factors, three intrinsic characteristics have facilitated the diversification of cichlids. First, most cichlids exhibit an efficient form of brood care, maternal mouth brooding: females incubate their eggs and fry in their buccal cavity, which protects the offspring in densely packed fish communities (Sturmbauer et al. 2011). Second, the cichlid jaw has been regarded a key-innovative trait. Cichlids possess two sets of jaws: the oral jaw and the pharyngeal jaw. Both jaws have experienced extensive diversification allowing for adaptation to a wide array of feeding modes and food resources (Liem 1973). Third, long range dispersal is rare in many cichlid species. This extreme philopatry together with the mosaic like distribution of habitat allows for a high degree of microallopatry (Danley et al. 2000, Rico and Turner 2002).

The Lake Malawi Cichlids

Lake Malawi, with more than 700 species, is the most diverse of the three East African Great Lakes. Most phylogenetic studies suggest that the majority of these species are monophyletic, despite multiple cichlid invasions of the lake (Albertson et al. 1999, Salzburger & Meyer 2004, but see Joyce et al. 2011 for an alternative opinion). Species communities are very diverse and provide a natural laboratory for evolutionary studies (Ding et al. in review).

The large scale pattern of diversification in the lake has been identified and described as three stage model in Danley and Kocher (2001) and Streelman and Danley (2003). During the first two stages, splits into macro-habitat clades and trophically diverged genera, were likely driven by strong divergent natural selection. More recently, it has been shown that one of the major habitat clades, the deep-benthics, was the result of an ancient hybridization event during low water levels (Genner & Turner 2012). In contrast to the first two stages, during the third stage of the radiation strong sexual selection likely resulted in the differentiation in male signaling phenotypes and reproductive behaviors leading to the isolation of closely related sympatric species (Kidd et al. 2006, Danley 2011, Danley et al. 2012b).

Spatial and Temporal Population Structure—The Effects of Drift and Migration

While most species are reproductively isolated, allopatric populations of the same species still exchange genes (Danley et al. 2000, Won et al. 2005). However, several studies have shown that cichlid populations phenotypically and genetically diverge despite ongoing gene flow (Arnegard et al. 1999, Markert et al. 1999, Danley et al. 2000, Streelman et al. 2007, Genner et al. 2010, Danley et al. 2012b). Whether this is the result

of drift or selection is not fully understood, but some authors have suggested that divergence with gene flow is the result of slight changes in the selective optima of local populations rather than drift (Danley et al. 2000). Further studies investigating the driving forces of such divergence are needed to provide insight into what has led to the extensive speciation of Malawi cichlids.

While spatial population structure has frequently been studied in a variety of Lake Malawi cichlids (e.g. Markert et al. 1999, Albertson et al. 1999, Danley et al. 2000), the temporal structuring of populations has been largely neglected. Highly dynamic temporal population structure might be a possible explanation for the rapid evolution observed in Lake Malawi cichlids. The few studies addressing temporal dynamics in population structure of cichlid fishes have revealed conflicting results. A study indirectly suggested high temporal dynamics of population genetic structure in a Lake Malawi cichlid: Streelman et al. (2004) showed genetic diversification of populations after an introduction event within a 20 year period. Another study of population structure in a riverine cichlid showed strong between-year variations of genetic structure (Crispo and Chapman 2010). In contrast, a study of between-year variations in a lacustrine cichlid did not reveal any temporal variation (Genner et al. 2010). The conflicting results of these studies are likely the result of the differences in time scales at which they were performed, habitat characteristics (deep lake environment vs. dynamic stream environment) and strongly divergent ecology of the studied species. In order to understand how likely drift is to affect populations of Lake Malawi cichlids on ecological time scales more detailed comparative studies of microendemics and widespread species are needed.

Ecological Selection

Ecological selection has caused the divergence of macrohabitat clades and genera in the first two stages of the Lake Malawi cichlid radiation (Danley & Kocher 2001). In addition, studies of widespread cichlid species have identified the divergence in ecological characters at the population level (Streelman et al. 2007, Pauers 2011). Selection is often thought to drive this phenotypic divergence, yet, the influence of drift, the null model for evolutionary change, needs to be tested. Further, it is unclear if the observed phenotypic differences are due to heritable genetic changes or are a plastic response to local environmental pressures. Tests are needed to quantify the influence of genetic and plastic responses of phenotypes (Kerschbaumer et al. 2011 for a Lake Tanganyika example).

Evolutionary ecological studies, however, are rare for Lake Malawi due to the remote location and logistic problems. The few studies performed so far have found that cichlid diversity is linked to structural heterogeneity of the habitat (Parnell & Streelman 2011, Ding et al. in review) and while most abiotic factors are fairly consistent throughout the lake (Ding et al. in review), resource composition differs among locations (Reinthal 1990, Abdallah & Barton 2003, Higgins et al. 2003). Accordingly, a population level study has shown that divergence in body shape and trophic morphology can also arise within species (Streelman et al. 2007). Yet, closely related sympatric species are found to use similar resources at the same location (Genner et al. 1999, Martin & Genner 2009). However, the sympatric species differ in microhabitat use, territory size, and territory defense (Holzberg 1978, Danley 2011, Albertson 2008). In order to understand how species with superficially similar ecologies can co-exist without competitively

exclusion more detailed studies on ecological differentiation of closely related sympatric species are needed.

Sexual Selection

While the importance of natural selection on the diversification of genera has been demonstrated in the past (Fryer & Iles 1972, Danley & Kocher 2001, Albertson & Kocher 2006), it has been frequently suggested that sexual selection might be more important at the level of species and maybe populations of Lake Malawi cichlids (Dominey 1984, Seehausen 2000, Danley & Kocher 2001, Pauers 2011). Strong sexual selection on male sexual characters is likely due to a strongly skewed parental investment in the offspring. Within the rock-dwelling cichlids females disproportionately invest in their offspring and hence are the choosy gender: females produce large yolky eggs with are orally incubated for about 20 days after fertilization. During that time females do not eat. In contrast, males have to acquire and defend territories in order to mate, but only contribute gametes during reproduction. Hence, male-male competition for territories as well as female mate choice place strong selective pressures on male mating phenotypes (Dijkstra et al. 2005, 2006, Danley 2011).

Traditionally it has been thought that visual cues, mainly different color patterns and hue are the single most important character influencing female mate choice in the mbuna. A variety of experiments, both in the field and in the lab, were conducted to demonstrate that divergence in visual traits among species is sufficient to cause reproductive isolation via assortative mating (Seehausen and van Alphen 1998, van Oppen et al. 1998, Couldrige and Alexander 2002, Maan et al. 2010). Several studies investigated within-species among-population color polymorphisms and showed that

similar color morphs are due to convergence and not monophyly (Arnegard et al. 1999, Smith and Kornfield 2002, Allender et al. 2003, Pauers 2011). These studies suggest that the evolution of color differentiation is highly dynamic and likely results from the strong effect of sexual selection at the species level. Yet, little is known on how selection acts on coloration of isolated populations within species. As gene flow is still present between populations purifying selection would be expected to keep coloration phenotypes at a species specific optimum. If closely related congeners occur in sympatry reinforcement or character displacement might act to emphasize species differences. Empirical studies are needed to test these assumptions.

Hybridization and Diversification

More recently an additional process has been recognized to be an important contributor in generating biodiversity in cichlids: hybridization (Salzburger et al. 2002, Streelman et al. 2004, Koblmüller et al. 2007, Seehausen 2004, 2013, Parsons et al. 2011, Joyce et al. 2011, Genner & Turner 2012, Schwarzer et al. 2012). Hybridization can generate new genetic and phenotypic variation on which selection can act (Seehausen 2004, Lucek et al. 2010). More and more examples accumulate which demonstrate the diversifying effect of hybridization in cichlids: A study by Genner & Turner (2012) for example demonstrated that a species rich macro-habitat clade of cichlids likely resulted from an ancient hybridization event. On the level of species it has been demonstrated that hybridization of species can lead to strong transgressive segregation in a variety of phenotypes (Albertson & Kocher 2005, Stelkens et al. 2009, Parsons et al. 2011) and that the amount of transgressive segregation found in a cross is correlated to the genetic distance between species (Stelkens et al. 2009, Stelkens & Seehausen 2009). Most of

these studies, however, investigated crosses between fairly distantly related species belonging to different genera. Therefore, relatively little is known on the effects of hybridization of closely related congenerics.

Goals

Historically, selection is considered the driving force in the divergence of cichlids in specific and many adaptive radiations in general; yet, the null hypothesis of divergence by drift has rarely been examined. Despite a large body of literature we are still at the beginning of understanding which forces have driven the radiation of East African cichlids. In this work I focus on the forces driving the divergence of populations of *Maylandia zebra* and closely related pairs of species with different coloration in the genus *Maylandia*.

In chapter two I review the importance of time series in evolutionary and conservation biology. This serves as an introduction to chapter three where I present a study using the temporal sampling approach in two species of rock-dwelling cichlids to estimate their effective population size and temporal stability to evaluate the importance drift might play in the evolution of these species. The estimates of population sizes can further be used to make suggestions for species conservation. The fourth chapter addresses the relative importance of drift and selection for the evolution of a sexual (color pattern) and two ecological characters (body shape and length) in ten populations of *Maylandia zebra*. The fifth chapter focuses on the potential role of correlational selection on body shape and coloration in replicate pairs of sympatric barred and non-barred species of *Maylandia*. Specifically, I discuss how co-existence in sympatry is facilitated if species diverge in sexual and ecological traits. The chapter raises the question if

differences observed in the field are genetic or plastic. This is addressed in chapter six where I use a common garden experiment and hybrid crosses between a barred and a non-barred species to study the plastic and genetic components of body shape, estimate the degree of transgressive segregation and discuss the importance of hybridization in the evolution of the cichlid radiations.

CHAPTER TWO

The Relevance of Time Series in Molecular Ecology and Conservation Biology

Introduction

Populations can undergo strong fluctuations in population size from one generation to the next. Usually these changes have a relatively small effect on the genetic composition of the population. Yet, depending on the population size and the cause of the fluctuation a significant change in genetic composition of species can occur in a single generation. Therefore, information about a population collected from single points in time often yield an incomplete picture of the historical and ongoing biological processes influencing populations (Crispo & Chapman, 2009; Husemann *et al*., 2012). Especially when the impacts of natural or anthropogenic events, which took place at a specific time point, are studied, only samples taken before and after the event may provide the information needed to understand the effects on the population.

For example, many studies have documented the genetic impact of population bottlenecks as a result of overharvesting and habitat destruction (e.g. Hauser *et al*., 2002; Frankham 2005), the differentiation of populations in response to limited connectivity and restricted gene flow (e.g. Danley *et al*., 2000; Husemann *et al*., 2012), the impact of introduced invasive species on native species (e.g. Ficetola *et al*., 2008; Ray *et al*., 2012) and species responses to climate change (e.g. Ayre & Hughes, 2004; Chaloupka *et al*., 2008). However, all of these referenced studies draw conclusions based on data collected from a single point. While studies have suggested that single year samplings are

sufficient to provide a good estimate of the genetic composition of a population (see Gomaa *et al*., 2011), multiple sampling points can be used to empirically explore the demographic history of populations and document the persistence of population structures (e.g. in naturally fragmented habitats). Temporal population genetic studies can quantify the effects of natural and anthropogenic factors on populations generate robust estimates of their effective population sizes. In addition, temporal designs can to test for the loss of genetic diversity, or to show an increase in population differentiation as a result of increasing population isolation and/or lower effective population sizes (e.g. Harper *et al*. 2006, Crispo & Chapman 2009). The vast amount of biological material stored in museum collections in combination with advanced DNA sequencing techniques make it possible to study the intraspecific effects of environmental and population changes over time (Luikart *et al*., 2003). Furthermore, the combination of whole genome scans using Next Generation Sequencing (NGS) and temporal population samplings allows the identification of changes in selective pressures over generations (Nielsen *et al*., 2009; Allendorf *et al*., 2010; Hohenlohe *et al*., 2010; Stapley *et al*., 2010; Gompert *et al*., 2010). While studies focusing on population responses to environmental conditions have often been carried out *ex-situ* in experimental setups with artificial selective regimes (Ball *et al*., 2000, Bijlsma *et al*., 2000, Kristensen *et al*., 2008, Reed *et al*., 2002), the use of time series may allow researchers to study the impacts of anthropogenic disturbance and large scale changes of environmental conditions (e.g. climate, nitrogen loads) to understand whether taxa or local populations have the genetic diversity required to adapt to future environmental changes within relatively short time periods.

In this review, we explore the potential biological materials, marker systems and associated limitations for time series studies. We discuss the advantages of analysing time series in molecular ecology and conservation biology (i) to estimate effective population size and the impact of random genetic drift, (ii) to explore the demographic history of populations (e.g. population fluctuations and population bottlenecks), and (iii) to study the impacts of changed habitat features and the relevance of habitat histories on inter- and intraspecific levels of the genetic structure.

Suitability of Samples and Markers

Analyses of populations sampled at multiple points in time are becoming increasingly relevant in the field of modern population biology, especially in population genetics and population genomics (e.g. Wandeler *et al*., 2007; Nielson & Hansen, 2008; Gomaa *et al*., 2011). However, museum collections rarely harbour sufficient numbers of suitable samples. Such samples need to be collected from the same generation and the same location to avoid unaccounted structure in the data. In addition, the sample needs to be stored in a manner such that the DNA is preserved and easily extracted (Nielson & Hansen, 2008). Therefore, studies need to be planned according to the historical material available. Thus this historical material should be located, DNA should be isolated and markers tested. Testing the markers is particularly important since even in cases where vast amounts of samples are available genotyping may not be possible. Contemporary sampling should only be performed after these preparations have been accomplished.

The DNA quality of historic samples strongly depends on the way organisms were collected, conserved and stored as well as on the age of these samples (Dean & Ballard, 2001). Some chemicals such as formalin and ethyl-acetate can degrade DNA

(Dillon *et al*., 1996; Schander & Halanych, 2003). Storing samples frozen or in high concentrations of ethanol can be costly and time-consuming and their storage requires space and organization. While many universities and museums have established cryobanks (Lermen *et al*., 2009), most samples, especially those interesting for studies addressing the genetic diversity before environmental changes took place, have been prepared conventionally by pinning and air-drying or curing (animals), or mounted on paper (in case of annual herbaceous plant species). These methods generally preserve samples in a way that allows additional morphometric comparisons which continue to be important ancillary information for population genetic data sets (see below). Given the persistent importance of morphological methods and the museum policies on destructive sampling, minimally invasive and non-destructive DNA extraction protocols have been developed in order to minimize the damage to specimens while maximizing the DNA yield (Mundy *et al*., 1997; Gilbert *et al*., 2007, Tagliavia *et al*., 2011).

Despite the development of new DNA isolation techniques, the variable and often highly degradation DNA from historic samples limits the choice of genetic markers available for population genetic analyses. Some genetic techniques, especially methods based on protein and RNA molecules, require very specific sample storage conditions and are generally not applicable to historic samples. In contrast, DNA can be well preserved in old biological materials and small fragments of DNA have successfully been amplified from samples 100 to 100,000 years old (e.g. Hofreiter *et al.*, 2002; Strange *et al*., 2009; Hoeck *et al*., 2010). However, the unsuitable storage of samples often leads to degradation so that only small fragments of DNA are available for analyses. This again limits the genetic markers that can be reliably genotyped for population analyses

(Wandeler *et al*., 2007; Nielson & Hansen, 2008). Methods which are based on the analyses of fragment polymorphisms, such as RAPD, RFLP and AFLP, are generally less suitable since highly degraded DNA can lead to misleading results due to homoplasy in these types of markers. The sequencing of larger genes or gene fragments can also be very difficult for degraded samples. In degraded samples, rarely are large genes left intact for sequencing and often multiple primer pairs have to be used to obtain the complete targeted fragment. For sequence analyses mitochondrial DNA (mtDNA) is preferred over nuclear DNA (nuDNA) because mtDNA occurs at higher copy numbers. Yet, mtDNA sequences are unable to detect reticulate events. As a result, nuclear markers are also desired in population studies. Given the fragmented nature of degraded nuDNA samples, methods targeting small fragments of DNA are preferred. Here, generally two types of markers are most commonly used: microsatellites and Single Nucleotide Polymorphisms (SNPs). These high-resolution markers are suitable for detecting genetic changes over short temporal and restricted spatial scales and are therefore frequently used in landscape and conservation genetics (Selkoe & Toonen, 2006). Little is known about the evolution of microsatellites, making it difficult to employ suitable evolutionary models during the analyses (Ellegren, 2004). In contrast, the bi-allelic SNPs are less variable compared to microsatellites, and represent the most common type of polymorphism in the genome. Due to its simplicity and broad range of applications, this marker system has become increasingly popular recently and is considered the marker of the future in population genetics (Morin *et al*., 2004).

The recent development of Next Generation Sequencing (NGS) techniques facilitates the discovery and typing of large numbers of genes and gene fragments (Baird

et al., 2008; Davey & Blaxter, 2010; Davey *et al*., 2011; Ekblom & Galindo, 2011; Hohenlohe *et al*., 2011). Most NGS methods employed in population genomics use genome reduction techniques including restriction digests and fragment size selection (Hohenlohe *et al*., 2011; Gompert *et al*., 2010). This allows for the simultaneous selective amplification of the desired number of loci across a number of specimens. Dozens of individuals can be barcoded and then pooled within runs on a NGS platform. However, often a two stage approach is employed where NGS methods are used only for marker discovery (e.g. Seeb *et al*., 2011, De Pristo *et al*., 2011, Davey *et al*., 2011). In a second step suitable markers are chosen from the large sets discovered by NGS and are then genotyped using qPCR-based approaches, high resolution melting (HRM) curve methods or Sanger sequencing. In the future, the costs of SNP genotyping in non-model organisms will decline further and will allow for the detection of large numbers of SNPs to detect population genetic structures even over many generations (Allendorf *et al*., 2010).

However, due to the above-mentioned limitations in sample availability one has to be aware that temporal studies will always be limited to relatively few species for which suitable material is available. Such flagship species will have to serve as representatives for other organisms with similar ecology and life histories.

Effective Population Size and Random Genetic Drift

In populations which are geographically isolated and where gene flow is low or lacking, genetic drift can be one of the main evolutionary forces driving divergence. The effect of genetic drift is largely determined by the effective population size (N_e). Small populations are generally more vulnerable to random processes than large populations. In contrast, drift is thought to play a minor role in large and interconnected populations

because random processes are balanced by a large number of copies of alleles being present in the gene pool; further, high allelic variability can be maintained by continuous gene flow among neighboring populations. Populations with a large effective population size can rapidly response to selection while smaller populations may lack the genetic diversity necessary to respond to similar selective pressures. However, generally it is difficult to determine the force having the strongest impact on a species or a population.

One way of inferring the relative impact of drift versus selection is to quantify the effective population size (N_e) (Franklin & Frankham, 1998). This is because the relative contribution of drift or selection is a function of N_e and the selection coefficient. In general, neutral alleles are governed by drift and non-neutral variants by selection. However, since selection is more effective in large populations where random events (drift) have a smaller impact, there is a threshold at which non-neutral alleles become effectively neutral and thus governed by drift. This threshold is given by the equation $4N_es = 1$, where N_e is the effective population size and s is the selection coefficient. Thus, in very small populations, even deleterious alleles suffering from large selection coefficients may become fixed and reduce the fitness of the population. This, in turn, may lead to a mutational meltdown in which the population size continues to decline leading to the fixation of more deleterious alleles which then causes a further decline in population size and so on (Lynch *et al*., 1993).

Temporal studies are effective at estimating N_e by examining the change in allele frequencies through time. Stable allele frequencies reflect a large N_e; fluctuating allele frequencies on the other hand indicate a small N_e. While not being the only method, time series can be used to estimate N_e and yield the most robust results (Barker, 2011).

Several studies have shown that drift can result in significant population divergence over different time frames. A study by Hoeck *et al*. (2010) showed that the degree of population divergence through drift strongly depends on habitat size which correlated with effective population size. The smaller a population, the higher the degree of population divergence it experienced over the study period of ~200 generations. Similarly strong genetic drift was shown by Harper *et al*. (2006) for a butterfly species. Here, elevated drift was the result of a dramatic population reduction due to the decline in the butterfly's trophic resource. Other studies yielded similar results of temporal instability and significant changes of allele frequencies over different time scales (e.g. Heath *et al*., 2002; Morris *et al*., 2002; Breinholt *et al*., 2009; Griffith *et al*., 2009). Analyses of the Yellowstone grizzly bear *Ursus arctos* show a strong decline in genetic diversity between 1912 and 1981 in addition to reduced individual viability. The decline in the population's overall fitness may be a consequence of this population's genetic impoverishment (Miller & Waits 2003).

However, changes in population structure need not always result in changes in the intraspecific genetic variability, but instead can lead to strong changes in genetic differentiation. For example the genetic differentiation of *Erysimum cheiranthoides,* an annual plant common on stony river banks, increased threefold from 2005 to 2007, while the genetic diversity remained fairly constant through the years (Honnay *et al*., 2009). High gene flow rates between the 16 studied populations and the relatively recent origin of the metapopulation structure may explain why recurrent extinction and colonization have not caused a decrease of genetic diversity. The authors argue that persistent seed

banks play an important role inin both maintaining the genetic diversity and in structuring the population after the moderate flooding event of 2007.

In contrast to these examples in which genetic diversity decreased and/or genetic differentiation increased from past to present, other studies showed `temporal stability´ of population structures without significant shifts over time. Within a brown trout (*Salmo trutta)* population, genetic diversity and population structure experienced little change over a period of 20 years (Palm *et al*., 2003). This example is in agreement with data obtained for the Leopard Frog, *Rana pipiens,* where five populations were studied over 22-30 years (equivalent to 11-15 generations). The data indicate stable and very large effective population sizes and temporal stability of its genetic structure (Hoffman *et al*., 2004). These studies highlight that extant genetic structuring is strongly affected by past population dynamics which has a direct impact on genetic drift and gene flow.

Effect of Population Bottlenecks

It is well known that demographic changes have the strongest impact on a population's genetic diversity (Frankham *et al*. 2004) and temporal molecular analyses represent powerful tools to analyse these changes. Of the many ways in which a population can experience a demographic change, population bottlenecks produce to the greatest genetic change due to genetic drift. In this case population sizes are drastically reduced and only a subset of the original diversity of a population is maintained. A textbook example was provided by Bouzat *et al*. (1998) who studied a population of the greater prairie chicken, *Tympanuchus cupido,* over a period of 30 years. The authors detected a large proportion of alleles which were exclusively found in historical samples, but were absent in recently collected wild individuals. The authors coined the term `ghost

alleles´ for these variants which are exclusively found in old sampling material, but have vanished in contemporary populations (Bouzat *et al*., 1998). They argue that these ghost alleles disappeared due to strong population fluctuations and subsequent population bottlenecks. The reduction of genetic diversity in this example was significantly correlated with a decline in the population size (due to habitat loss) and finally caused a decrease in individual fitness (see also Hansson & Westerberg, 2002; Reed & Frankham, 2004; Leimu *et al*., 2006). Similar trends of reductions in genetic diversity over generations in the wake of habitat transformation and associated reduced population sizes have also been found in other animal and plant species (e.g. Harper *et al*. 2006; Groombrige *et al*., 2000). When comparing seedlings of the highly endangered tree endemic to the Seychelles, *Vateriopsis seychellarum,* collected in pre- and post-fragmentation populations, the genetic data show a severe decline in genetic diversity together with an increase in genetic differentiation. The authors explain these effects as a consequence of the rapid reduction in the number of trees and low gene flow rates among local populations (Finger *et al*., 2012).

The detection of ghost alleles in historical samples collected in a population need not necessarily imply a reduction in the total number of alleles or past population bottlenecks. The Violet Copper butterfly *Lycaena helle* has been geographically restricted to small and isolated habitats at higher elevations in the Middle Mountains of Central Europe since the postglacial warming. A comparison of its recent genetic diversity with individuals collected 15 years ago identified strong shifts in allele frequencies, the vanishing of many alleles (i.e. the existence of ghost alleles), but a relatively stable count in the total number of alleles over generations – despite its existence in rather small and

isolated populations (Habel *et al*., 2011) (see Fig. 2.1). This is consistent with a study by Harper *et al*. (2006) which showed large changes in allele frequencies but a stable number of alleles in a butterfly species over a time frame of about 100 years. These observations are consistent with a population in drift-mutation-migration equilibrium. Populations experiencing such equilibrium lose alleles due to drift at the same rate that migration and mutation introduce new neutral alleles to the population. In general, such populations are considered to be fairly stable over the studied time frame (Piry *et al*., 1999).

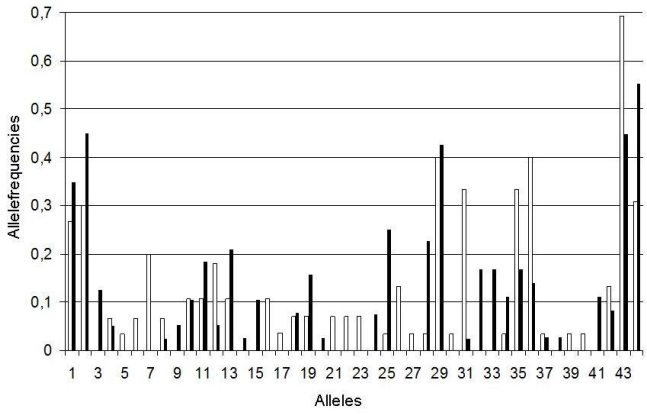

Figure 2.1: Allele frequency shift for one microsatellite locus analysed in the butterfly *Lycaena helle* of one population (France, Massif Central, Mareuge) for the year 1991 (white) and 2006 (black). A clear shift in alleles and their frequencies were detected whereas the total number of alleles remained similar (migration-drift-mutation equilibrium) (34 in the year 1991 and 31 for the year 2006). Data taken from Habel *et al*. (2011).

Likewise, the loss of genetic diversity does not always indicate a recent population bottleneck. A geographically restricted relict population of the Red Apollo butterfly, *Parnassius apollo,* in the Mosel valley of western Germany was almost completely monomorphic at six microsatellite loci (Habel *et al*., 2009) that were polymorphic in French populations of the same species (Meglecz et al. 2004). The

genetic impoverishment of the Mosel valley populations was hypothesized to be the result of a severe population collapse during the 1960s as a result of indiscriminate insecticide spraying. However, these microsatellites, though polymorphic in other populations, were monomorphic in the Mosel valley population and this monomorphism was stable across two temporal collections. Samples collected before (1890-1960) and after (1960-today) both showed this lack of diversity at these microsatellite loci. This indicates that *P. apollo* was already genetically impoverished before the population collapsed. Similar population genetic stability despite small and isolated populations can be found in the endangered Seychelles endemic *Medusagyne oppositifolia* which naturally occurs only on inselberg habitats (granitic outcrops). Here, despite fragmentation, the species was able (at least in its largest population) to maintain a high genetic diversity when comparing adult trees with progeny (Finger *et al*., 2011).

In summary we can delineate three different population genetic processes: (i) the loss of genetic diversity over time due to genetic drift in isolated populations, (ii) migration-mutation-drift equilibrium, in which the loss of alleles is offset by the introduction of new alleles through migration and mutation, and (iii) the persistence of intraspecific diversity despite severe population bottlenecks as a consequence of long-term isolation.

The use of historical samples to detect, quantify and interpret potential effects of recent population bottlenecks, however, must be carried out with caution. Conclusions are only valid if historical sample sizes are representative (which is often not the case) and co-dominant markers can be reliably genotyped. This is best highlighted by pointing out the consequences of a bottleneck on heterozygosity and allelic diversity.

Heterozygosity is often quite insensitive to bottlenecks and even a population decline to two individuals will only lead to a loss of heterozygosity of $1/(2N_e)$ = 25% in one generation (see Allendorf, 1986). In contrast, two individuals can only possess a maximum of four different alleles. This makes allelic diversity a better parameter for bottleneck detection. However, the effect of a bottleneck on allelic diversity depends on the total number of alleles found in a population and their frequencies, whereas the rate at which heterozygosity declines is always $1/(2N_e)$ regardless of the initial heterozygosity (Allendorf & Luikart, 2007). Accurate estimates of the number of alleles and their frequencies strongly depend on sample size, which is why a representative sampling of historical populations is critical.

The Relevance of Habitat Histories – Habitat Persistence Versus Habitat Transformation

Apart from population fluctuations due to environmental stochasticity and subsequent population bottlenecks, additional extrinsic forces play an important role in shaping the genetic makeup of populations. The fragmentation of formerly interconnected habitats typically increases the population structure within the species and fractures its genetic cohesiveness both of which often have a negative impact on the species (Zachos *et al*., 2007; Sharma *et al*., 2011). Some species, however, appear to be more tolerant to habitat fragmentation than others (e.g. Valqui *et al*., 2010). It is challenging to explain such contrasting responses to changes in habitat structures, which in turn makes it difficult to develop appropriate conservation strategies for species whose habitats are currently being destroyed. So far, only a few studies have explored the importance of population demographic histories in understanding a species' vulnerability to the negative consequences of habitat fragmentation (but see Angeloni *et al.*, 2011;

Leimu & Mutikainen, 2005). Empirical studies using samples collected before and after fragmentation events for a variety of taxa are necessary to understand what affects the vulnerability of species to fragmentation.

Current research attempts to explain contrasting responses to changes of the environmental conditions such as habitat fragmentation (Leimu *et al*., 2006; Angeloni *et al*., 2011; Finger *et al*., 2012). For example, species that have historically existed in large, interconnected population networks may have been able to exchange genes among local habitat patches over short distances. Rapid and drastic environmental changes that disrupt these metapopulations may result in the sudden reduction or loss of gene flow, population differentiation, loss of genetic diversity through increased drift and may finally result in inbreeding depression. One example comes from tropical East Africa where forests with different habitat histories experience different levels of habitat degradation (Habel & Zachos, 2012). The Chyulu Hills in southern Kenya are a naturally fragmented forest-meadow mosaic, while the neighbouring Taita Hills have suffered severe human-induced habitat destruction over the past few decades. The Mountain White-eye *Zosterops poliogaster* inhabits both of these habitats and genetic data from this species reflect the divergent habitat histories of these now similar habitat structures. The Chyulu Hills population, collected in 1938 and 2011, maintained its genetic diversity and no genetic differentiation was detected in contemporary subpopulations. In contrast, the Taita Hills population sampled at different time points over the past 20 years shows a strong increase in genetic differentiation among local subpopulations (Habel *et al*., 2013). Together these findings demonstrate that fragmented habitat conditions (Chyulu Hills) do not necessarily have a negative impact on the intraspecific genetic diversity of a species *per se*, whereas

fast transformations from interconnected to highly fragmented ecosystems (Taita Hills) may severely affect the biota living there. Such a sudden collapse of formerly intact habitat and metapopulation networks and the associated transition from (near-)panmixia to situations of reduced gene flow often have a negative impact on the maintenance of genetic variability, and result in strong deviations from Hardy-Weinberg and high inbreeding coefficients (Kadlec *et al*., 2010; Konvicka *et al*., 2010). In the ground beetle *Carabus violaceus,* for example, recent habitat fragmentation due to road construction caused the split of a local population into two subgroups, resulting in strong genetic differentiation (Keller & Largiadèr, 2003; Keller *et al*., 2005). The rapid development of genetically differentiated populations was also observed in the riverine cichlid fish, *Pseudocrenilabrus multicolor victoria.* Data suggests that this species experienced dramatic changes of its intraspecific structure across only a few years. While a clear isolation by distance pattern was detected in the first year, the pattern was wiped out in the following sampling year. The authors suggested that this rapid change was the result of a severe flood between the sampling years (Crispo & Chapman, 2009) (Fig 2.2). Rapid genetic responses were also found in large mammals with comparatively long generation intervals. Fickel *et al*. (2012) found that only 20 years (or roughly three generations) after the fall of the iron curtain, panmixia was re-established in red deer from the Bavarian-Bohemian forest ecosystem. In contrast to these examples, other studies indicate a genetic time lag for organisms living in changing habitat structures, as shown for the ground beetle *Carabus auronitens*, which today occurs in interconnected forest habitats, but still displays strong patterns of differentiation in as a result of its past pattern of distribution (Drees *et al*., 2008).

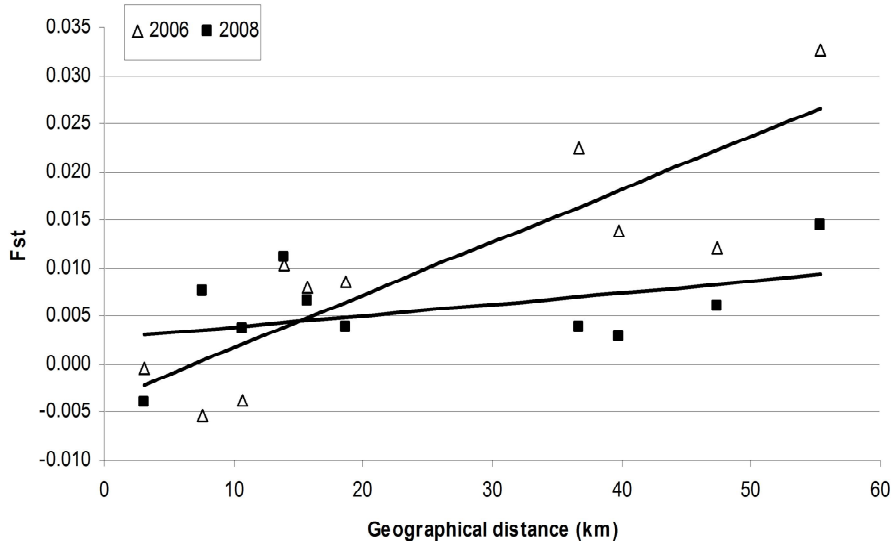

Figure 2.2: An example for temporal instability in genetic structure; the same populations of the riverine cichlid *Pseudocrenilabrus multicolor victoriae* were sampled before and after a flooding event. The original isolation-by-distance pattern found before the flooding event was eradicated after the event (from Crispo & Chapman 2009).

The above-mentioned examples highlight how the ecology of species and their demographic history affects their genetic structure. Owing to intrinsic species-specific requirements (such as microclimate or host specificity), some species may adapt to persisting in small and isolated populations and may have existed in such systems with low connectivity over long time periods. The species´ historical distribution and population structure may have a substantial influence on its response to recent environmental changes (which may be less negative in generalist and genetically diverse taxa). Inbreeding depression, for example, may be lower in populations that have been small for a long time and may have consequently purged deleterious alleles, whereas a recent reduction in population size may cause stronger inbreeding depression (Lande & Schemske, 1985; Keller & Waller, 2002). Still, there seem to be controversies as to whether rare, widespread, endemic or non-endemic species will be more prone to

negative genetic consequences of habitat fragmentation. Common species and large populations were found to be as, or even more, susceptible to the loss of genetic diversity through habitat fragmentation as rare species and small populations (Angeloni *et al.*, 2011; Honnay & Jacquemyn, 2007). It becomes apparent that in order to analyse the effects of environmental changes on populations, researchers have to consider both the past (e.g. pre-fragmentation, pre-bottleneck) and recent (e.g. post-fragmentation, post-bottleneck) population structure of the organism. Only analyses including multiple temporal samples of a population will be able to empirically disentangle recent rapid effects from past long-term processes.

CHAPTER THREE

Only Time Will Show—Analyses of Census and Effective Population Sizes and the Demographic History of Populations of the Lake Malawi Cichlid Genus *Maylandia*

Introduction

A central goal in evolutionary biology is to identify the forces leading to population divergence and speciation. On ecological time scales the interplay of three forces defines population structure: migration, drift and selection. Divergent selection and drift promote the fixation of alleles and therefore lead to population differentiation while migration homogenizes populations and prevents such differentiation. The relative importance of each force depends on the geographic setting, local selective pressures, intrinsic species specific characteristics, historic and current demography of populations and their effective population size. The latter, defined as the size of an ideal population that experiences evolutionary forces in a manner equivalent to the observed population (Fisher 1930), is one of the most important variables in population and conservation genetics (Luikart et al. 2010).

The effective population size (N_e) determines how vulnerable a population is to stochastic forces and determines the rate of change as a result of drift (Vucetich et al. 1997, Palstra & Ruzzante 2008, Charlesworth 2009). Therefore, estimating N_e in natural populations is a first step in understanding the forces leading population differentiation (Charlesworth 2009). Drift has a stronger effect in small, isolated populations with low migration rates, whereas populations with large N_e are less affected by stochastic events. While drift can be counterbalanced even by low levels of gene flow, selection can

overcome the effect of migration and lead to population divergence even in the face of steady gene flow (Nosil 2008).

Another important application of N_e is to estimate the viability of a population. Effective population sizes of 500-5000 are considered necessary for species to maintain their evolutionary potential (Lynch & Lande 1997, Franklin & Frankham 1998, Traill et al. 2007). Therefore, estimating N_e can be helpful in making informed species management and conservation decisions.

However, estimating N_e in natural populations is not trivial as many of the most reliable approaches require genetic samples from populations taken at multiple points in time (Pollak 1983, Waples 1989, Jorde & Ryman, 1995, William & Slatkin 1999, Wang & Whitlock 2003, Barker 2011). Such samples can be rare especially in organisms in which samples were not specifically stored for future genetic analysis (Wandeler et al. 2007, Nielson & Hansen 2008, Habel et al. in press). Despite these problems generating temporal genetic data can be rewarding and help to answer long posed evolutionary questions. One such question is the relative impact of drift in the diversification of rapid species radiations (Markert et al. 1999, Kornfield & Smith 2000).

The East African cichlids are the most diverse vertebrate radiation known. More than 2000 species have evolved in the three East African Great Lakes: Lakes Victoria, Tanganyika and Malawi (Danley et al. 2012). While Lake Tanganyika's cichlids invaded the lake and diversified over the past 6-12 million years, the cichlid species flocks in Lakes Malawi and Victoria are of recent origin and the current species assemblages have evolved most likely within the last million years (Danley et al. 2012). Lake Malawi harbors more than 700 species. The diversification of the cichlid fauna in the lake has

been strongly influenced by the geographic and climatic history of the region (Genner et al. 2010, Danley et al. 2012, Aguilée et al. 2011, 2013) as well as the shore and basin morphology (e.g. Danley et al. 2000, Genner et al. 2010). The patchy distribution of habitat types together with the low dispersal rate of most species has led to high degree of microendemism in the rock-dwelling cichlids. Most microendemics are found at a single, often small, location in the lake suggesting that population sizes are small and migration cannot mitigate the effect of drift. Therefore, stochastic processes may play a profound role in the evolution of these species. Few species are found at multiple locations throughout the lake and previous population genetic studies on these cosmopolitan species have demonstrated that these populations are connected by low levels of gene flow (Markert et al. 1999, Arnegard et al. 1999, Danley et al. 2000). Therefore, these populations are expected to have larger effective sizes relative to microendemics and may be in drift-migration equilibrium.

In this study we test if the microendemic species *M. benetos* has a small effective population size and is influenced by drift. We further compare this microendemic to several populations of the widespread species *M. zebra* which is expected to have a larger effective population size and to be in migration-drift equilibrium. We use mitochondrial DNA sequences and nuclear microsatellites for temporal population samples to estimate contemporary effective population sizes, infer the demographic histories and test for migration-drift and mutation drift equilibia. We further generated estimates of census size for the populations to estimate N_e/N_c ratios, another important estimator in conservation biology (Frankham 1995, Luikart et al. 2010).

Materials and Methods

Study Species and Locations

The genus *Maylandia* is among the most diverse genera of rock-dwelling cichlids in the lake with 31 described species (Cicotto et al. 2011, Stauffer et al. 2013). *Maylandia zebra* is among the few species that can be found at nearly every rocky habitat throughout the lake. In contrast, most *Maylandia* species are microendemics and occur only at a single location in the lake. One of these is *Maylandia benetos* which is endemic to Mazinzi Reef where it co-occurs with two other *Maylandia* species, one of which is *M. zebra* (Danley 2011, Husemann et al. in review). Mazinzi Reef (-14.142818, 34.965091) is a small submerged reef about 10,000 m^2 in size. The sandy habitat around the reef and the relatively large distance to the next stretch of rocky shoreline make it fairly isolated relative to our other sampling locations, Illala Gap (-14.000119, 34.848361) and Harbour Island (-14.06894, 34.92965). Illala Gap and Harbour Island are well connected sites at the edge and in the center of large stretches of rocky habitat (Fig. 3.1, Danley et al. 2000).

Specimen Collection

In 1996, 2010 and 2012, we collected *Maylandia zebra* at 3 locations (Mazinzi Reef, Harbour Island, Illala Gap, Fig. 3.1, Table 3.1). *Maylandia benetos* was collected at Mazinzi Reef in the same years. Specimens were caught using SCUBA and gill nets. Fish were fin clipped in the field and fin tissue was either dried, stored in ethanol or preserved in a DMSO storage solution (20 % DMSO, 0.25 M EDTA, saturated with NaCl, pH = 7.5) until further processing.

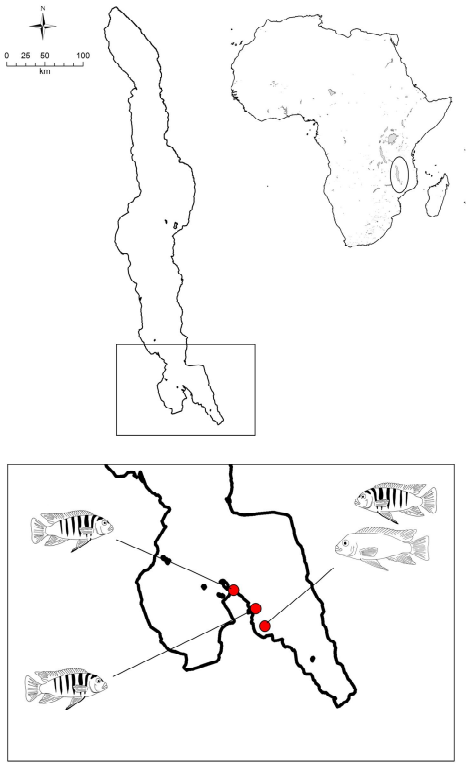

Figure 3.1. Sampling map.

Table 3.1. Sampling list indicating the sample numbers for each year and each population including the number of individuals sampled for microsatellites, D-loop sequences and the respective Genbank accession numbers.

Year	Location	# individuals microsatellites	# D-Loop sequences	Genbank accession numbers
1996	Harbour	31	29	KC960349-KC960377
	Illala	30	29	KC960320-KC960348
	Mazinzi	29	28	KC960407-KC960434
	benetos	30	29	KC960378-KC960406
	Total	120	115	
2010	Harbour	30	25	KC960277-KC960301
	Illala	30	18	KC960302-KC960319
	Mazinzi	30	25	KC208879 - KC208904
	benetos	29	29	KC208850 - KC208878
	Total	119	97	
2012	Harbour	29	27	KC960250-KC960276
	Illala	30	26	KC960172-KC960197
	Mazinzi	26	25	KC960225-KC960249
	benetos	27	27	KC960198-KC960224
	Total	112	105	
all		351	317	

Mitochondrial DNA

DNA was isolated using the Qiagen DNeasy blood and tissue kit (Qiagen, Hilden, Germany) following the manufacturers protocol for tissue samples. We amplified a 442 bp fragment of the mitochondrial D-Loop for 143 specimens (Table 1, 2) using the forward primer HapThr-2: 5' CCTACTCCCAAAGCTAGGATC '3 and the reverse primer Fish12s: 5' TGCGGAGACTTGCATGTGTAAG '3 (Joyce et al. 2005). PCR was performed using the following setup: 12.2 µl of diH_2O, 2 µl of 10x PCR buffer (reaction concentration 1x), 1.6 µl of dNTP mixture (0.2 µM each, Thermo Scientific), 0.2 µl of DyNAzyme™ DNA Polymerase (1.2 U, Thermo Scientific), 1 µl of each primer (0.5 µM, Integrated DNA Technologies, Coralville, IA, USA) and 2 µl of DNA template (either pure extract, 1:10, 1:50 dilution) for a total volume of 20 µl. Amplification conditions were as follows: 94ºC for 3 min., followed by 30 cycles of 94ºC for a 1 min. denaturation, 58ºC for a 1 min. annealing and 72ºC for a 2 min. elongation, with a final elongation step at 72ºC for 10 min. 10 µl PCR product were purified using 4 µl ExoSAP-IT enzyme mix (Affymetrix, Santa Clara, CA, USA). The purified products were sequenced at the Sequencing Facility at Yale University. All sequences were deposited in Genbank (Accession # are given in Table 3.1).

We aligned our sequences using Geneious v. 6.0.3 (Drummond *et al*., 2011); base calls were checked by eye. General statistics of sequence variation were calculated with DnaSP v.5.10 (Librado & Rozas, 2009). Estimates of genetic differentiation among species, populations and temporal samples were calculated as Φ_{ST} with Arlequin v. 3.5.1.2 (Excoffier and Lischer 2010) and were tested for significance using 100 permutations. Analysis of molecular variance (AMOVA) was used to partition genetic

variance on three levels: between populations, between temporal samples and within temporal samples.

We estimated the demographic history of each population independently using Bayesian Skyline analysis (Drummond et al. 2005). As strict clocks are considered appropriate for intraspecific data (Hein et al. 2005), we used a published substitution rate of 0.0324 changes per site per million years (SE 0.0139) for the D-Loop in cichlids which has been independently estimated by Genner et al. (2010) and Koblmüller et al. (2011). The HKY + I substitution model was determined as most suitable for the data with JModeltest 2.1.1 (Posada 2008). The temporal samples for each population were pooled due to limited genetic differentiation (see AMOVA results). Each analysis was run for 100 million steps sampling every 10,000 steps under default settings. The output from BEAST was subsequently analyzed in Tracer v. 1.5 (Rambaut & Drummond, 2009) from which the data were exported into Excel and displayed. To further explore the demographic histories of the populations, mismatch distributions were calculated in Arlequin.

Microsatellites

A total of 13 microsatellite loci were genotyped for 351 individuals in this study (Table 3.1, 3.3, Appendix 3.1). The loci were chosen to be on different linkage groups (Albertson et al. 2003) to ensure they were physically unlinked. Microsatellites were amplified using standard PCR procedures and labeled using the technique described by Schuelke (2000). We used the fluorescent dyes VIC, 6-FAM and PET for fragment visualization. The master mix for each reaction consisted of: 14.7 µl diH_2O, 2.0 µl 10x buffer (Thermo Scientific), 1.6 ul dNTPs (Thermo Scientific, 0.2 µM each), 0.1 µl

forward primer + M-13, 0.4 µl reverse primer (0.5 µM, Integrated DNA Technologies, Appendix 3.1), 0.2 µl Taq (1.2 U, DyNAzyme, Thermo Scientific), and 1 µl of template. The following amplification conditions were used: initial denaturation for 5 min. at 94°C, followed by 30 cycles of 30s at 94°C, 45s at 56-61°C (depending on locus, Appendix 3.1), and 45s at 72°C, followed by 8 cycles of 30s at 94°C, 45s at 53°C, and 45s at 72°C for M-13 binding, concluding with a 10 min. elongation at 72°C. PCR products were vizualized on 1 % agarose gels stained with gel red (Biotium, Inc. Hayward, CA). Three markers (each 1 µl) with different labels were subsequently pooled and denatured with 7 µl Hi-Di Formamid (Applied Biosystems). Fragment analysis was performed at the Sequencing Facility at Yale University using LIZ-500 as size standard. Genotypes were determined from chromatograms manually using the microsatellite plug-in in Geneious. Micro-Checker (Van Oosterhout et al. 2004) was used to test our data for genotyping errors due to stutter and null alleles. Create 1.37 (Coombs et al. 2008) and populations v. 1.2.32 (Langella 1999) were used to generate input files for downstream population genetic analyses.

Standard population genetic analyses were performed in Arlequin v. 3.5.1.2 (Excoffier and Lischer 2010), Genepop on the Web (Raymond & Rousset, 1995), and FSTAT v. 2.9.32 (Goudet, 1995). We estimated the number of alleles across all loci per population, the mean number of alleles per locus, the observed heterozygosity (H_o), the expected heterozygosity (H_e), and allelic richness (AR). We tested for deviations from Hardy-Weinberg Equilibrium (HWE), and Linkage Disequilibrium (LD) and applied a Bonferroni correction to account for multiple simultaneous tests. A Hierarchical Analysis of Molecular Variance (AMOVA, Excoffier et al. 1992) was performed to partition the

molecular variance on three levels: the four sampling locations were defined as highest hierarchical level, followed by the temporal samples within populations and last level partitioned diversity within populations. The genetic divergence among locations and temporal samples was calculated as R_{ST} based on the sum of squared size differences and tested for significance using 100 permutations (Slatkin 1995).

Equilibrium Testing

The software Bottleneck v. 1.2.02 (Cornuet & Luikart 1996, Piry et al. 1999) was used to test if populations are in mutation-drift equilibrium using the most recent temporal sample (2012). As the mutation model is generally not well understood in microsatellites we tested our data under all three available models (Infinite Allele Model (IAM), Stepwise Mutation Model (SMM), Two-Phase Model (TPM)). The SMM is the simplest model and assumes mutations happen in single repeat steps (Ohta & Kimura 1973). The IAM assumes mutations are random and equivalent alleles are identical by descent (Kimura & Crow 1964). The TPM model was developed more recently and allows both, single and multi-step mutations and assigns different probabilities to each type of mutation (DiRienzo et al. 1994). As this model implements stepwise mutations in addition to multi-step mutations it may conform to microsatellite data more appropriate than the other models (Neff et al. 1999). Wilcoxon signed-ranks statistics were used to test if observed heterozygosity exceeded the expectation under mutation-drift equilibrium.

To test for migration-drift equilibrium the program 2mod v. 0.2 was used (Beaumont 2000). The program estimates the relative likelihoods of drift alone vs. drift-migration to identify populations subjected to genetic drift and migration as opposed to those only affected by drift. For this temporal samples were pooled as temporal variation

was very limited (see AMOVA). We performed 2 separate runs each with 100,000 iterations. The first run included all 3 populations of *M. zebra*, whereas the second run contained *M. benetos*. The first 10 % of iterations were discarded as burn-in.

Effective Population Size

We employed five different approaches to estimate the effective population size of our study populations. Of the five, one method uses a single temporal sample and provides estimates based on gametic disequilibrium between alleles from different neutral loci as implemented in the LD option in N_eEstimator (Ovenden et al. 2007) The other four methods use multiple temporal cohorts, but are based on different principles: TM3 is based on Bayesian statistics and assumes closed populations sampled at two time points not far apart from each other in time (1996, 2012) so that mutation has a negligible effect on the observed gene frequencies. TempoFS (Jorde & Ryman 2007) estimates genetic drift between temporally spaced samples using the F_s measure of allele frequency change. MLN_e (Wang & Whitlock 2003) uses a maximum likelihood approach to estimate drift between temporally spaced populations, whereas N_eEstimator employs a temporal method using moments-based F-Statistics (Pollak 1983). As all approaches have their advantages and disadvantages (Schwartz et al. 1998) we calculated the harmonic means of all temporal estimates as our final estimate for N_e for each population (Waples 2005, Johnstone et al. 2012).

Transect Data and Census Size

We estimated the density and census size for all populations. Habitat sizes for our study locations were estimated previously (Danley et al. 2000). To estimate the density at

each location three transects of 25 m each were established at different depth (5m, 10m, 12-15m). Each transect was observed once by P.D.D.; every individual of *M. zebra* and *M. benetos* within 1 m to each side of transect was recorded. We used two methods to calculate density and census size. First we simply calculated the mean of sightings across all transects within a location and estimated the density by dividing with the area observed (50 m^2). Using the density we calculated the census size of each population by interpolating to the habitat size (Danley et al. 2000). In addition we used the program Distance v. 6.0 (Thomas et al. 2009, 2010). This program uses distance sampling data and habitat size to estimate the density and census size of populations. The specific distance for each individual to the transect line could not be observed as the fish were moving. Therefore, the perpendicular distance was set to 1 m for each individual as this was the maximum observation range.

Results

Mitochondrial DNA

A total of 317 individuals were sequenced (18-29 per population, Table 3.2). A total of 30 haplotypes were recovered with a total of 25 segregating sites. The total haplotype diversity was 0.820. The largest number of haplotypes founding any populations and temporal cohort was 9 for *M. zebra* from Harbour Island sampled in 1996. Harbour Island was generally the most diverse population and had the highest haplotype number in every sampling year (6-9 haplotypes) and a total of 14 haplotypes across years. The least diverse population was *M. benetos* from Mazinzi with 3 haplotypes found in each sampling year and a total number of 5 haplotypes across all

years. *Maylandia zebra* from Mazinzi Reef and the Illala Gap were intermediate with 4-5 haplotypes in each year and a total of 9 and 10 haplotypes across years, respectively.

Table 3.2. Sequence summary statistics for the D-Loop (sequence length 695 bp).

Year	Location	# sequences	# haplotypes	Segregating sites	Haplotype diversity	K (average # differences)	Pi (nucleotide diversity)
1996	Habour	29	9	13	0.83990	4.10099	0.00587
	Illala	29	4	5	0.25369	1.33498	0.00191
	Mazinzi	28	5	6	0.68254	1.90212	0.00273
	M. benetos	29	3	2	0.24631	0.30542	0.00045
	Total	115	19	16			
2010	Habour	25	8	10	0.66667	4.12667	0.00597
	Illala	18	5	7	0.57516	1.49708	0.00215
	Mazinzi	25	4	4	0.63667	1.65333	0.00238
	M. benetos	29	3	3	0.06897	0.20690	0.00030
	Total	97	16	15			
2012	Habour	27	6	8	0.77208	3.41026	0.00493
	Illala	26	4	3	0.45231	0.49538	0.00071
	Mazinzi	25	5	5	0.47000	0.90667	0.00131
	M. benetos	27	3	2	0.33048	0.34758	0.00050
	Total	105	15	17			
Complete dataset		317	30	25	0.820	3.418	0.00495

AMOVA showed that most of the variation is found among the four populations (57.31 %, Table 3.4a). Most of the remaining variance was found within temporal samples (39.29 %). Very little variance was explained by differences between temporal samples (3.4 %). This was corroborated by Φ_{ST} estimates which were high and significant between and low and mostly non-significant within populations (Table 5).

The Bayesian Skyline analyses yielded different populations sizes and demographic histories for each of the populations (Fig. 3.2). The largest contemporary population size was found for *M. zebra* from Harbour Island. The population appears to have started to expand about 6 kya and still is expanding. Similarly the *M. zebra* population from Mazinzi Reef expanded during that time, yet is much smaller than the

one found at Harbour Island. The expansion of the Illala Gap population started more recently (4 kya), but the population reached less than a third of the size of Harbour Island. The population of *M. benetos* from Mazinzi Reef appears to be of very recent origin (<1 kya) and has by far the smallest effective population size compared to *M. zebra* populations.

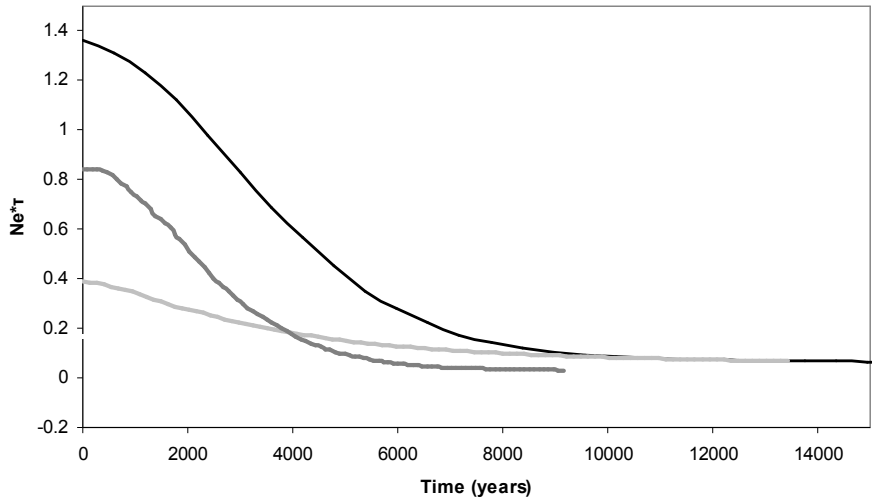

Figure 3.2. Bayesian Skyline Plots. Bayesian skyline plots derived from mtDNA control region haplotypes, where the y-axis represents a product of effective population size (N_e) and generation time (τ, in millions of years); black – *M. zebra* Harbour, dark grey – *M. zebra* Mazinzi, light grey – *M. zebra* Illala, interrupted – *M. benetos* Mazinzi.

Mismatch distributions were used to further explore the demographic history of populations. A separate mismatch distribution was calculated for each temporal sample (Appendix 3.2). The mismatch distributions of temporal samples of each population were largely consistent: *M. zebra* from Harbour Island and Mazinzi Reef show two strong peaks, whereas *M. zebra* from the Illala Gap and *M. benetos* from Mazinzi only exhibit a single distinct peak.

Microsatellites

We amplified 13 microsatellite loci for a total of 351 individuals. Micro-Checker did not find evidence for allele drop out or genotyping error in any sample; yet, it suggested the presence of null alleles for some markers at temporal samples of populations. No marker, however, consistently showed a null allele problem across multiple populations (Table 3). Global analyses of LD with Genepop suggested no significant linkage existed for any of the loci. HWE was confirmed for most loci across populations and temporal samples after Bonferroni correction.

The only exception was UNH2135 for *M. benetos* and *M. zebra* in 1996 at Mazinzi Reef. The deviation from HWE was caused by an access of homozygotes at these loci. Population subdivision is an unlikely reason for this excess in these populations as they exist at a small submerged reef; therefore, Wahlund effects can be excluded as an explanation for the heterozygote deficit. Null alleles are not a likely explanation as Micro-Checker did not identify null alleles for this locus in these populations. Inbreeding, however, might be a viable explanation for the heterozygote deficit (Galbusera et al. 2004), especially since Mazinzi Reef is the smallest studied location and *M. benetos* is a microendemic species.

The number of alleles found within temporal samples of populations varied between 186 for *M. benetos* in 1996 to 299 alleles detected for *M. zebra* from Harbour Island in 2010 (Table 3). The most diverse population was *M. zebra* from Harbour Island with a total of 415 alleles across all time points and an average number of 31.9 alleles per locus. The least diverse population was *M. zebra* from Mazinzi Reef with a total of 276 alleles and an average of 21.2 alleles per locus across all temporal samples.

Table 3.3: Summary statistics of microsatellite data (calculation of allelic richness was based on a minimal sample size of 20 individuals).

Population	Sampling year	# alleles	Average # alleles	Ho	He	Allelic richness*	Possible null alleles
Habour	1996	277	21.3 (SD 4.8)	0.936	0.937	18.212	UNH2135
	2010	299	23.0 (SD 5.4)	0.863	0.936	19.355	UNH2037, UNH362, UNH2204
	2012	275	21.2 (SD 5.9)	0.924	0.934	18.318	-
	All	415	31.9 (SD 9.3)	n.a.	n.a.	31.718	n.a.
Illala	1996	282	21.7 (SD 4.8)	0.909	0.938	18.468	UNH2135, UNH362
	2010	250	19.2 (SD 3.4)	0.897	0.928	16.619	UNH2155, UNH2139
	2012	263	20.2 (SD 4.0)	0.898	0.931	17.518	UNH2112
	All	399	30.7 (SD 6.4)	n.a.	n.a.	29.479	n.a.
Mazinzi	1996	211	16.2 (SD 3.9)	0.909	0.897	14.229	-
	2010	193	14.8 (SD 4.2)	0.863	0.882	13.150	-
	2012	184	14.2 (SD 4.8)	0.897	0.881	12.999	UNH362
	All	276	21.2 (SD 6.7)	n.a.	n.a.	20.641	n.a.
M. benetos	1996	206	15.8 (SD 3.9)	0.895	0.890	13.855	UNH2037, UNH2190
	2010	186	14.3 (SD 3.9)	0.870	0.884	12.899	UNH2166
	2012	193	14.8 (SD 4.4)	0.917	0.881	13.380	-
	All	287	22.1 (SD 6.4)	n.a.	n.a.	21.343	n.a.

Accordingly, allelic richness was highest for the Harbour Island population with an average of 31.718 across all loci and temporal samples. In comparison, *M. zebra* from Mazinzi had the lowest allelic richness with 20.641 (Table 3). Estimates of observed heterozygosity ranged from 0.863 for *M. zebra* from Harbour Island and Mazinzi Reef in 2010 and 0.936 for *M. zebra* from Harbour Island in 1996.

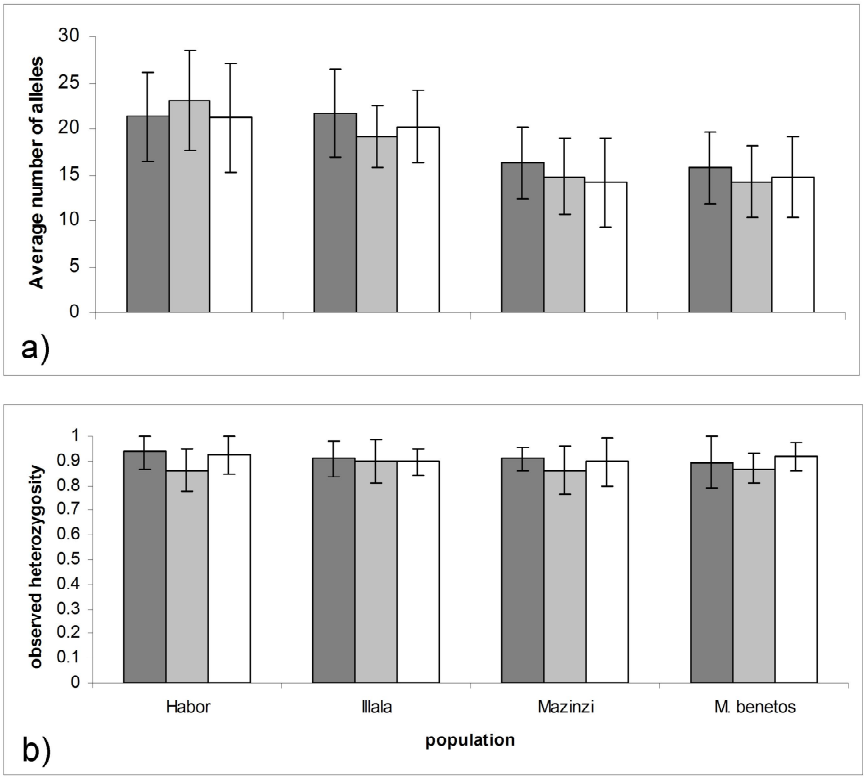

Figure 3.3. Comparison of genetic diversity estimates for the four studied populations across the three sampled time points: a) mean number of alleles per locus and b) observed heterozygosity (H_o); dark grey 1996, light grey 2010, white 2012, Tukey HSD tests showed that all comparisons between years within populations were not significant (α = 0.05) indicating similarity of the average number of alleles and heterozygosity within populations across all sampling years.

AMOVA detected the majority of genetic variance within temporal samples (83.09 %, Table 3.4b). Some variance was found between populations (16.35 %), whereas very little variance was explained by the different temporal samples within populations (0.56 %). Estimates of genetic divergence corroborate this finding with inter-population comparisons being mostly significant (except Harbour Island 2010 to all Illala Gap samples), whereas intra-population comparisons yielded very small and mostly non-significant results (Table 5).

Table 3.4a. AMOVA for mtDNA data; groups defined as locations/populations, temporal samples defined within groups.

Source of variation	d.f.	Sum of squares	Variance component	% variance	p-value
Among populations	3	329.720	1.34959 Va	57.31	<0.001
Among temporal samples within populations	8	24.233	0.08005 Vb	3.40	0.003
Within temporal samples	305	282.202	0.92525 Vc	39.29	<0.001
total	316	636.155	2.35490		

Table 3.4b. AMOVA of microsatellite data using the R_{ST} statistics; groups defined as locations/populations, temporal samples defined within groups.

Source of variation	d.f.	Sum of squares	Variance component	% variance	p-value
Among populations	3	321446.984	586.91248 Va	16.35	<0.001
Among temporal samples within populations	8	33332.226	20.25744 Vb	0.56	0.027
Within temporal samples	690	2057836.437	2982.37165 Vc	83.09	<0.001
total	701	2412615.647	3589.54157		

Table 3.5. Φ_{ST} D-Loop (below diagonal) and R_{st} from microsatellites above diagonal (Mb – *M. benetos* from Mazinzi Reef, MR – *M. zebra* from Mazinzi Reef, HI – *M. zebra* from Harbour Island, IG – *M. zebra* from Illala Gap, 12 – sampling date 2012, 10, sampling date 2010, 96 – sampling date 1996).

Pop.	Mb12	Mb10	Mb96	MR12	MR10	MR96	HI12	HI10	HI96	IG12	IG10	IG96
Mb12	0	0.010	0.012	**0.319**	**0.316**	**0.306**	**0.298**	**0.199**	**0.259**	**0.178**	**0.191**	**0.188**
Mb10	0.025	0	**0.024**	**0.291**	**0.285**	**0.270**	**0.248**	**0.139**	**0.214**	**0.121**	**0.137**	**0.126**
Mb96	0.009	0.040	0	**0.314**	**0.319**	**0.303**	**0.324**	**0.215**	**0.280**	**0.205**	**0.204**	**0.220**
MR12	**0.764**	**0.785**	**0.780**	0	-0.011	-0.012	**0.092**	**0.083**	**0.075**	**0.143**	**0.135**	**0.142**
MR10	**0.538**	**0.561**	**0.558**	**0.087**	0	-0.008	**0.087**	**0.069**	**0.060**	**0.122**	**0.112**	**0.120**
MR96	**0.414**	**0.433**	**0.433**	**0.177**	-0.015	0	**0.098**	**0.077**	**0.085**	**0.123**	**0.129**	**0.130**
HI12	**0.556**	**0.576**	**0.567**	**0.427**	**0.344**	**0.348**	0	**0.031**	0.019	**0.065**	**0.092**	**0.056**
HI10	**0.552**	**0.576**	**0.560**	**0.487**	**0.406**	**0.397**	**0.110**	0	0.006	-0.001	0.009	0.002
HI96	**0.541**	**0.560**	**0.552**	**0.399**	**0.324**	**0.333**	-0.024	0.107	0	**0.037**	**0.032**	**0.028**
IG12	**0.909**	**0.923**	**0.916**	**0.825**	**0.768**	**0.754**	**0.572**	**0.335**	**0.539**	0	**0.014**	0.005
IG10	**0.844**	**0.864**	**0.854**	**0.770**	**0.698**	**0.679**	**0.489**	**0.208**	**0.465**	**0.254**	0	0.01
IG96	**0.827**	**0.845**	**0.835**	**0.751**	**0.686**	**0.669**	**0.488**	**0.194**	**0.461**	**0.130**	0.061	0

Equilibrium Testing

Our two-tailed Wilcoxon tests for mutation-drift equilibrium revealed significant deviations from mutation-drift equilibrium for *M. zebra* from Mazinzi under the IAM (p

= 0.033) and SMM ($p < 0.001$) models but not under the TPM model ($p = 1.000$). The Illala Gap population tests were only signification under the IAM ($p < 0.001$) model (TPM $p = 0.340$, SMM $p = 0.455$). The Harbour Island population tests were significant under the IAM ($p < 0.001$) and SMM ($p = 0.048$) models but not TPM ($p = 0.068$). Tests of mutation-drift equilibrium for *M. benetos* suggested deviations from the equilibrium only for the SMM ($p = 0.002$) model, but not for the TPM ($p = 0.542$) and IAM ($p = 0.057$) models. As the TPM model incorporates aspects of IAM and SSM it can be considered to provide the most accurate estimate (Neff et al. 1999). None of our populations showed deviations from drift-mutation equilibrium for the TPM model.

The test for migration-drift equilibrium with 2mod indicated that the *M. zebra* populations are in migration-drift equilibrium (100 % support), whereas the signal was less clear for *M. benetos* in which the effects of drift were greater than migration and migration-drift disequilibrium better explained the data than the equilibrium model (57.7 %).

Effective Population Size

The estimates of N_e varied between populations and methods (Table 6). Estimates from single sample methods were generally lower and are considered less reliable when compared to temporal methods (Barker 2011); therefore we focus our discussion mostly on the temporal data. However, Bayesian Skyline analyses on mitochondrial data and N_e estimates on microsatellites generally showed the similar trends. *Maylandia zebra* from Harbour Island consistently has the largest estimates population size, indicated by both, the mitochondrial and microsatellite data (Fig. 3.2, Table 3.6, except MLN_e). In contrast, the microendemic *M. benetos* had the smallest effective population size for most methods

(except TM3). The estimates for *M. zebra* from the Illala Gap generally yielded very small population sizes and were similar to N_e estimates for *M. benetos*. The *M. zebra* population from Mazinzi Reef had relatively large population sizes which were intermediate between those of the Harbour Island and Illala Gap populations. The harmonic means across all temporal methods were very similar for *M. zebra* from Harbour Island (N_e = 1279.5, 95 % CI: 616.6-2862.4) and Mazinzi Reef (1213.1, 462.8-14006.6) and for *M. zebra* from Illala Gap (557.2, 336.6-1282.5) and *M. benetos* from Mazinzi Reef (510.4, 237.4-1357.4).

Census Size

Our estimates of census sizes calculated from three transects for each location indicate very different population sizes at each location. However, both estimates of census size (mean interpolation, distance) were highly similar (Table 6). The individual densities were fairly similar across all locations and ranged from 0.520 individuals / m^2 for the Illala Gap population to 0.881 individuals / m^2 for *M. zebra* at Mazinzi Reef. The Harbour Island population had a density of 0.775 individuals / m^2, whereas the density of *M. benetos* was relatively low with 0.533 individuals / m^2. The population census size estimates ranged from 5,333 individuals for *M. benetos* and 8,811 individuals for *M. zebra* from Mazinzi Reef up to 1,106,820 individuals for *M. zebra* at Illala Gap. The estimated census size for Harbour Island was intermediate with 129,113 individuals. The N_e/N_c ratios differed strongly as well and ranged from $1.38*10^{-1}$ for *M. zebra* from Mazinzi Reef to $5.03*10^{-4}$ for *M. zebra* from the Illala Gap (Table 6).

Discussion

In this study we used microsatellites and mitochondrial sequences of temporally sampled populations belonging two species of Lake Malawi cichlids to investigate their population demographic history and estimate contemporary effective population sizes. We used our genetic data to check for temporal fluctuations in allele or haplotype frequencies to determine the effect of drift on the populations and to test if populations are in drift-migration and drift-mutation equilibrium. We further used transect data and estimates of habitat size to calculate census sizes which were compared to the effective population sizes.

Despite the importance of effective population size as parameter in evolutionary and conservation biology, few studies have estimated this parameter in cichlids. The only study we are aware of was performed by Won et al. (2005) who estimated effective and ancestral population sizes from composite genetic loci in species of the genus *Tropheops*. Interestingly, the study by Won et al. (2005) investigated the population size of *Tropheops gracilior* at Harbour Island and estimated N_e to be 3,600. Other species and populations in their study had even larger with N_e ranging from 1,500 to 47,800 (Won et al. 2005).The estimates of N_e we obtained for *Maylandia zebra* are lower and range from 557 for the Illala Gap population to 1,280 for the Harbour Island population. These population sizes are relatively small and are at the lower bound of what is considered necessary to maintain the evolutionary potential of populations (Lynch & Lande 1997, Franklin & Frankham 1998, Traill et al. 2007).

Table 3.6. Habitat size (Danley et al. 2000), census and effective population sizes based on microsatellite data for all studied populations estimated using different approaches. For single time point methods the most recent year (2012) was used, given are N_e with 95 % CI.

Population	Habitat size [m^2]	Census size (N_c)	Census size (N_c)	N_e (LD, Ne Estimator)	N_e (Ne Estimator)	N_e (TM3)	N_e ($TempoF_s$)	N_e (MLNe)	Average of all temporal methods	Ratio N_e / N_c
Sampling	See Danley et al. 2000	transect	transect	Single time point, most recent	Multiple time points	Multiple time points	Multiple time points	Multiple time points		
Method	See Danley et al. 2000	Mean interpolation of density	Distance sampling	Linkage Disequilibrium	Moments based	Bayesian approach	Genetic drift among samples	Maximum Likelihood	Harmonic mean* across all temporal estimates	
Illala *M. zebra*	2128500	1106820	1106820	190.1 (142.7-281.3)	631.4 (386.3 – 1276.7)	795.3 (372.0 – 1255.7)	353 (196 – 1617)	665.0 (462.3 – 1087.3)	557.2 (336.6-1282.5)	$5.03*10^{-4}$
Harbour *M. zebra*	166700	129026	129113	209.1 (153.4 – 329.9)	1966.2 (749.8 – ∞)	2966.8 (0 - 100000)	981 (442 – ∞)	793.8 (524.3 – 1452.9)	1279.5 (616.6-2862.4)	$9.91*10^{-3}$
Mazinzi *M. zebra*	10000	8800	8811	105.9 (80.4-152.7)	1093.2 (481.6-34319.8)	1794.0 (0 – 31787.9)	965 (346 – ∞)	1268.3 (661.7 – 6511.2)	1213.1 (462.8-14006.6)	$1.38*10^{-1}$
Mazinzi *M. benetos*	10000	5340	5333	97.9 (76.9 – 133.0)	554 (324 – 1217.8)	996.7 (420-1936.7)	289 (111 – ∞)	638.9 (422.7 – 1147.1)	510.4 (237.4-1357.4)	$9.57*10^{-2}$

*only estimates larger than 0 and lower than ∞ included

Yet, the populations exceed a N_e of 500 above which drift is considered to be negligible in populations (Hartl & Clark 2007). However, the mean estimate of N_e of *M. benetos* equaled 510.4 suggesting that drift might play a role in the evolution of *M. benetos*.

Population Sizes and Demographic History

These findings are consistent with other measures of population demographies. Bayesian Skyline analyses indicated that all populations expanded within the last 10,000 years and estimates population sizes for the *M. zebra* populations from Mazinzi Reef and Harbour Island were larger in comparison to the Illala Gap population. Populations of *M. zebra* showed temporal stability in the numbers of alleles and heterozygosity values (Fig. 3.3) and appear to be in migration- and mutation-drift equilibria. Therefore, the effect of drift is offset by gains in genetic variation by migration and mutation. Under these equilibrium conditions, populations are diverging (Table 5) consistent with the findings of Danley et al. (2000) who interpreted this pattern as divergence with gene flow.

Despite equilibrium conditions, effective population sizes differed among subpopulations of *M. zebra*. The population from Illala Gap is much smaller and has a lower genetic diversity than the Mazinzi Reef or Harbour Island populations. This might be the result of an edge effect as the population is located at the end of a longer stretch of continuous suitable habitat. It is known that populations at the edge of a distribution often have smaller population sizes and suffer from the loss of genetic diversity as a result of the stepwise loss of alleles from the centre of the distribution. Further, edge populations are often characterized by stronger population fluctuations relative to populations in the center of a range (Eckert et al. 2008, Gassert et al. 2013). That the Illala Gap population

has experienced more fluctuations is supported by the mismatch distributions which indicated that the Illala Gap population underwent a recent population bottleneck or expansion (Appendix 3.2). A population bottleneck, however, was not confirmed by the analysis of microsatellite data. Bayesian Skyline analyses also indicated a recent, but minor, population expansion. Therefore, it appears that the population has expanded recently and has reached equilibrium despite its relatively small size.

In contrast to the widespread *M. zebra,* the microendemic *M. benetos* has a relatively small effective population size ($N_e \sim 510$). While this N_e estimate is similar to the Illala Gap population of *M. zebra*, *M. benetos* does not have any populations which could serve as source of diversity via migration. This is supported by the 2mod analysis which indicated deviations from migration-drift equilibrium and suggested that drift is an important driver in this population. The mismatch distribution for *M. benetos* was unimodal indicating either a sudden population expansion or a recent population bottleneck (Harpending 1994). Bayesian Skyline analysis suggested that *M. benetos* originated within the past 1,000 years. This is in line with recent low levels of the lake which completely exposed Mazinzi Reef completely (Owen et al. 1990, Danley et al. 2012) and suggests that the species has evolved at that location after the inundation of the lake.

Despite evidence for drift in the microendemic, overall fluctuations of allele and haplotype frequencies were limited in all populations (Fig. 3.3) and we saw very little temporal signal across the sampling time of 16 years. AMOVA revealed that temporal samples only explained 3.4 % of the variance for mitochondrial data and 0.56 % for the microsatellite data (Table 4a, b). The average number of allele frequencies and the

observed heterozygosity were similar at all time points for all populations (Fig. 3.3). Likewise, measures of population genetic divergence (R_{ST} and Φ_{ST}) were significant for almost all inter-population comparisons, yet, small and mostly non-significant for intra-population comparisons between years (Table 5). This indicates that drift probably is not strong enough to cause significant population fluctuations over the ecological time scales sampled in this study.

Our comparison of effective population size and census size indicated highly skewed ratios across populations ranging from $5.03*10^{-4}$ for the Illala Gap population to $1.38*10^{-1}$ for *M. zebra* from Mazinzi Reef (Table 6). These estimates differ by three orders of magnitudes. However, they are in the range of estimates found for other fishes; for example in the snapper *Pagurus auratus* N_e/N_c was estimated to be $1.8–2.8 \times 10^{-5}$ (Hauser et al. 2002), in the red drum *Sciaenops ocellatus* N_e/N_c ratios were between $1.0*10^{-3}$ and $7*10^{-4}$ (Turner et al. 2002), in the pike *Esox lucius* estimates ranged between $3*10^{-2}$ and $1.4*10^{-1}$ (Miller & Kapuscinski 1997), for the Rio Grande Silvery Minnow *Hybognathus amarus* the ratio was found to be $\sim 1*10^{-3}$ (Alò & Turner 2004), and in the flatfish *Pleuronectes platessa* the N_e/N_c was estimated at 2×10^{-5} (Hoarau et al. 2005). Our estimates cover the whole range of these other studies.

The N_e/N_c ratio can be useful to make informed management decisions. If this ratio is stable, N_e can be estimated from N_c and vice versa (Luikart et al. 2010). In our case, however, we see strong deviations of these ratios across populations. We see a strong decline of the N_e/N_c ratio with increasing N_c. This is a common pattern and has been attributed to a high variance in male reproductive success in populations with a large N_c (Pray et al. 1996, Ardren & Kapuscinski 2003, Ficetola et al. 2010). Since the

reproductive behavior of haplochromine cichlids, including nearly all Lake Malawi cichlids, results in the high variance of male reproductive success due to male-male competition and female mate choice (Hert 1989, McKaye et al. 1990), we expect the decline of N_e/N_c ratio with increasing N_c. In order to understand the full extend in which N_e/N_c ratios are altered by male reproductive success future studies need to estimate both parameters in additional cichlid species.

Implications for Conservation

In addition to their evolutionary importance, the effective population size and census size are important parameters in conservation biology and can help to make informed management decisions (Luikart et al. 2010, Palstra & Fraser 2012). Despite their stunning diversity, few studies have focused on cichlid conservation. Yet several examples of large scale species extinctions as result of overfishing (Sturmbauer 2008, Sturmbauer et al. 2011), the introduction of introduced predators (Kaufman 1992, Gophen et al. 1993, Awiti 2011), species translocations (Genner et al. 2006, Zidana et al. 2009), and environmental destruction and pollution (Seehausen et al. 1997), all indicate that active cichlid management plans are needed (e.g. Sturmbauer 2008, Sturmbauer et al. 2011). One of the few conservation plans suggested the creation of micro-scale protected areas as means of protecting cichlid diversity (Sturmbauer 2008). We support this suggestion and further suggest that microendemic species, which as indicated in this study may maintain low effective population sizes that approach the minimum viable population size limit of 500 individuals, require special protection. As these populations might not have the evolutionary potential to cope with strong environmental stress their protection is crucial, especially since most microendemics are represented by single

populations. Future work should focus on generating N_e and N_c estimates for more microendemics in order to identify which species are in particular need of protection.

CHAPTER FOUR

A Comparison of Neutral Genetic and Phenotypic Divergence in Populations of the Lake Malawi Cichlid *Maylandia zebra*

Introduction

To understand rapid phenotypic changes, three evolutionary forces need to be quantified: selection, drift and gene flow. In general, selection is considered the driving force in most natural diversification processes as even small differences in the local selective regimes can lead to phenotypic changes. Yet, when population sizes are small stochastic sampling processes can lead to phenotypic change through genetic drift (Lande 1976, 1992, Spitze 1993, Hoeck et al. 2010, Whitlock 2000). However, low rates of gene flow will prevent divergence by drift (Slatkin 1987), whereas even slight differences in the selective environments can overcome the balancing effect of migration (Dobzhansky 1970, Danley et al. 2000, Gavrilets 2004). In natural populations it can be difficult to distinguish between the diversifying effects of selection and drift (Barton 1996, O'Hara 2005).

Comparisons of quantitative phenotypic data and neutral genetic markers can help to disentangle the influence of drift and selection on phenotypic divergence (McKay & Latta 2002, Leinonen et al. 2006, 2008, 2013, Sæther et al. 2007, Chenoweth & Blows 2008, Magalhaes et al. 2009, Hangartner et al. 2011). The influence of drift on the differentiation of phenotypes can be estimated from the divergence of neutral genetic markers as estimated by F_{ST} (Whitlock & Guillaume 2009). The value of F_{ST}, the estimate of population differentiation by drift, can then be compared to the differentiation

of phenotypes expressed as Q_{ST} (Spitze 1993), the differentiation in quantitative traits (Spitze 1993), or P_{ST}, the phenotypic alternative for Q_{ST} (Sæther et al. 2007, Raeymaekers et al. 2007). The F_{ST}/P_{ST} comparison assumes that the divergence in the quantitative traits is similar to the divergence in neutral genetic loci if the studied phenotypic traits evolve neutrally and have an additive genetic basis (Wright 1951, Lande 1992). If the phenotypic divergence is greater or less than neutral genetic divergence some form of selection can be implicated (Leinonen et al. 2006, 2008, 2013, Sæther et al. 2007, Whitlock & Guillaume 2009).

The cichlids of Lake Malawi belong to the most diverse vertebrate radiation known (Kocher 2004). More than 700 species have evolved in the last one million years from a single common ancestor (Albertson et al. 1999, Danley et al. 2012). Most species in the lake are microendemics found only at a single location in the lake (Stauffer et al. 1997, 2013, Allender et al. 2003). Only a few species are found at multiple locations, one of which is *Maylandia zebra*. Across their distributions, populations of such widespread species diverge genetically and phenotypically (Arnegard et al. 1999, Markert et al. 1999, Danley et al. 2000, Smith & Kornfield 2002, Allender et al. 2003, Streelman et al. 2007, Genner et al. 2010, Pauers 2011). As populations are fairly isolated and population sizes of many populations are relatively small the potential for divergence by drift might be high (Won et al. 2005, Husemann et al. in prep). However, low rates of migration persist at least on small and intermediate geographic scales (Danley et al. 2000). As even such low rates of migration should balance the effects of drift, selection has been invoked to be the driving force of such local differentiation processes (Danley et al. 2000, Streelman et al. 2007, Salzburger 2009, Pauers 2011).

In this study we investigate genetic and phenotypic differentiation in ten populations of the rock-dwelling cichlid *Maylandia zebra*. We use 10 microsatellites and mitochondrial sequences to generate estimates of neutral divergence and to test for selection in two ecological traits, standard length and body shape, and a sexual trait, standardized melanophore counts. A previous study has demonstrated variation in the body shape of local populations of *M. zebra* and suggested that body shape and coloration might be under correlational selection (Husemann et al. in review). Accordingly it could be expected that similar signatures of selection might be found for both traits. Alternatively, both traits could be uncoupled and their divergence driven by separate evolutionary forces.

Materials and Methods

Sampling

In 2010 and 2012, we collected *Maylandia zebra* at 10 locations throughout Lake Malawi: Mazinzi Reef, Harbour Island, Illala Gap, Mumbo Island, Nkhata Bay, Chiofu Bay, Domwe Island, Boadzulu Island, Otter Point, and Thumbi West Island (Fig. 4.1, Table 4.1). Specimens were caught using SCUBA and gill nets. Fish were photographed with a Canon Eos 540d and standard length was measured. Individuals were fin clipped and fin tissue was either dried or stored in a storage solution (20 % DMSO, 0.25 M EDTA, saturated with NaCl, pH = 7.5) until further processing. A scale of each fish was taken and transferred to aggregating solution (see below). The fish were released at their original location.

Table 4.1. Sampling list of specimens collected for this study; given are the numbers of individuals sampled for melanophore counts, and geometrics morphometrics on body shape, individuals genotyped for 10 microsatellite loci and the mitochondrial D-Loop with their respective Genbank accession numbers.

Population ID	Population name	# individuals melanophores	# individuals body shape/ standard length	# individuals microsatellites	# individuals mtDNA	Genbank accession numbers
1	Nkhata Bay	11	25	28	27	KF366222-KF366248
2	Chiofu Bay	16	25	29	24	KF366170-KF366193
3	Boadzulu Island	15	22	27	21	KC208919-KC208932
4	Mumbo Island	7	22	28	28	KF366194-KF366221
5	Otter Point	36	49	27	23	KC208951-KC208973
6	Thumbi West Island	15	27	27	23	KF366147-KF366169
7	Domwe Island	16	20	29	27	KF366120-KF366146
8	Illala Gap	20	41	30	37	KC960302-KC960319. KC960172-KC960197
9	Harbor Island	13	21	31	47	KC960277-KC960301, KC960250-KC960276
10	Mazinzi Reef	22	40	30	45	KC208879-KC208904, KC960225-KC960249
	Total	171	292	286	302	

Melanophore Counts

To quantify melanistic coloration we used standardized counts of melanophores on scales. For this scales were sampled from the first complete bar right above the lateral line (Danley 2001, Fig. 4.2). The scales were first incubated in a K^+-rich aggregating solution (modified from Vokey and Burton 1998) to condense the melanophores. Subsequently, scales were fixed in formalin and then transferred to Phosphor-buffered saline (PBS). Scales were photographed with a Nikon SMZ1500 under 30x magnification. All melanophores in a 0.25 mm^2 area were counted.

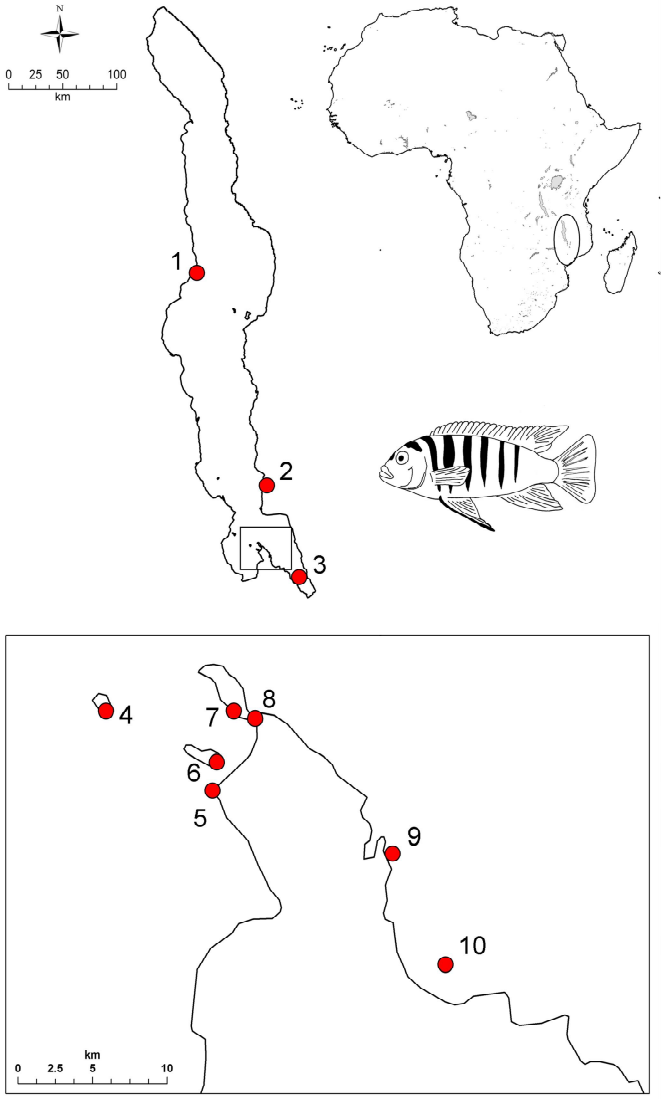

Figure 4.1. Sampling map. The population numbers correspond to those found in Table 4.1 (1 – Nkhata Bay, 2 – Chiofu Bay, 3 – Boadzulu Island, 4- Mumbo Island, 5 – Otter Point, 6 – Thumbi West Island, 7 – Domwe Island, 8 – Illala Gap, 9 – Harbour Island, 10 – Mazinzi Reef).

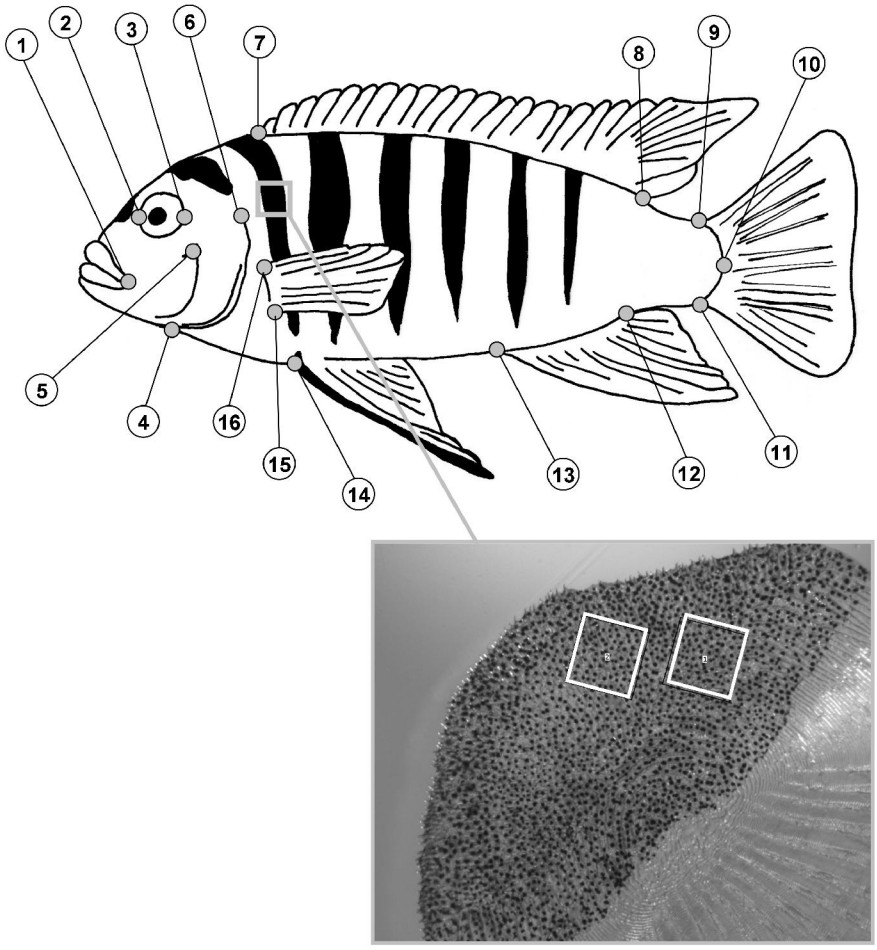

Figure 4.2. The 16 landmarks analyzed in this study for populations of *M. zebra*: 1) most posterior point of the lips, 2) anterior edge of the eye, 3) posterior edge of the eye, 4) ventral tip of cleithrum, 5) dorsal end of pre-opercular groove, 6) dorsal origin of operculum, 7) anterior insertion of dorsal fin, 8) posterior insertion of dorsal fin, 9) upper insertion of caudal fin, 10) midpoint of the origin of caudal fin, 11) lower insertion of caudal fin, 12) posterior insertion of anal fin, 13) anterior insertion of anal fin, 14) anterior insertion of pelvic fin, 15) ventral insertion of pectoral fin, and 16) dorsal insertion of pelvic fin. Location of scale sampled is shown as well as an example of a scale and the 0.25 mm^2 melanophores were counted in.

Geometric Morphometric Analyses of Body Shape

We quantified body shape variation in the ten *Maylandia zebra* populations using geometric morphometric analyses (Adams *et al*., 2004). Lateral pictures of individual fish were imported into tpsDig v.2.16 (Rohlf, 2006). On each picture 16 homologous

landmarks were identified and scored (Fig. 4.2). Landmark coordinates were aligned using least-square superimposition as implemented in the program tpsRelw (Rohlf, 2007). Based on the aligned coordinates, we calculated a weight matrix, which consists of the partial warp scores with uniform components for each individual. To reduce data dimensionality, we subjected the weight matrix to a principal component analysis (PC) based on the covariance matrix of the landmarks. Standard length and sex were included in the analyses as covariates to control for multivariate allometry. PC axes scores were then used as dependent variables in a multivariate analysis of covariance (MANCOVA). Assumptions of multivariate normal error and homogeneity of variances and covariances were met for all analyses performed. *F*-values were approximated using Wilks' lambda and effect strengths by use of partial eta squared (η_p^2). We included population identity as independent variable, and used the standard length and sex as a covariate to control for multivariate allometry. To visualize variation between populations, we calculated divergence scores for each individual based on population divergence vectors as defined by Langerhans (2009). Individual scores were then used as independent variables in tpsRegression (Rohlf 2005) to generate thin-plate spline deformation grids highlighting shape differences among populations (Zelditch et al. 2004). Unless otherwise stated, all statistical analyses were performed using SPSS 20 (IBM Inc.).

Mitochondrial DNA

DNA was isolated using the Qiagen DNEasy blood and tissue kit (Qiagen, Hilden, Germany) following the manufacturers protocol for tissue samples. We amplified a 646 bp fragment of the mitochondrial Control Region for 332 specimens using the forward primer HapThr-2: 5' CCTACTCCCAAAGCTAGGATC '3 and the reverse primer

Fish12s: 5' TGCGGAGACTTGCATGTGTAAG '3 (Joyce et al. 2005). PCR was performed using the following setup: 12.2 µl of diH_2O, 2 µl of 10x PCR buffer (reaction concentration 1x), 1.6 µl of dNTP mixture (0.2 µM each, Finnzymes, Vantaa, Finland), 0.2 µl of DyNAzyme™ DNA Polymerase (1.2 U, Finnzymes, Vantaa, Finland), 1 µl of each primer (0.5 µM, Integrated DNA technologies, Coralville, IA, USA) and 2 µl of DNA template (either pure extract, 1:10, 1:50 dilution) for a total volume of 20 µl. Amplification conditions were as follows: 94ºC for 3 min., followed by 30 cycles of 1 min. denaturation at 94ºC, 1 min. annealing at 58ºC and 2 min. elongation at 72ºC, with a final elongation step at 72ºC for 10 min. 10 µl PCR product were purified using 4 µl ExoSAP-IT enzyme mix (Affymetrix, Santa Clara, CA, USA). The purified products were sequenced at the Sequencing Facility at Yale University. All sequences were deposited in Genbank (Accession # are given in Table 4.1). Additional sequences were taken from precious studies (Husemann et al. in review, Husemann et al. in prep).

We aligned the sequences using Geneious v. 6.0.3 (Drummond *et al*., 2011); base calls were checked by eye. General statistics of sequence variation were calculated with DnaSP v.5.10 (Librado & Rozas, 2009). Estimates of genetic differentiation between populations were calculated as Φ_{st} with Arlequin v. 3.5.1.2 (Excoffier & Lischer 2010) using the Tamura-Nei substitution model and were tested for significance using 100 permutations. We tested for isolation by distance using the IbD web server v. 3.25 (Jensen et al. 2005).

We estimated the demographic history of each population independently using Bayesian Skyline analysis (Drummond et al. 2005). As strict clocks are considered appropriate for intraspecific data (Hein et al. 2005), we used a strict clock with a

published rate of 0.0324 changes per site per million years (SE 0.0139) for D-Loop substitution rates in cichlids (Genner et al. 2010, Koblmüller et al. 2011). The HKY + I substitution model was determined as most suitable model by JModeltest 2.1.1 (Posada 2008). Each analysis was run for 100 million steps (expect for the population from Mumbo Island which only converged after running 200 million iterations) sampling every 10,000 steps under default settings. The output from BEAST was subsequently analyzed in Tracer v. 1.5 (Rambaut & Drummond, 2009) from which the data were exported into Microsoft Excel and displayed.

Microsatellites

We chose an initial set of 14 microsatellite loci for our analyses (Appendix 3.1). The loci were chosen to be on different linkage groups (Albertson et al. 2003) to ensure markers are physically unlinked. Microsatellites were amplified using standard PCR procedures and labeled using the technique proposed by Schuelke (2000). We used the fluorescent dyes VIC, 6-FAM and PET for fragment visualization. The master mix for each reaction consisted of: 14.7 µl diH_20, 2.0 µl 10x buffer (Thermo Scientific), 1.6 ul dNTPs (Thermo Scientific , 0.2 µM each), 0.1 µl forward primer + M-13, 0.4 µl reverse primer (Appendix 3.1), 0.2 µl Taq (1.2 U, DyNAzyme, Thermo Scientific), and 1 µl of template. The following amplification conditions were used: Initial denaturation for 5 min at 94°C, followed by 30 cycles of 30 s at 94°C, 45 s at 56-61°C (depending on locus, Appendix 3.1), and 45 s at 72°C, followed by 8 cycles of 30 s at 94°C, 45 s at 53°C, and 45s at 72°C for M-13 binding, followed by a final elongation for 10 min at 72°C. PCR products were visualized on 1 % agarose gels stained with gel red (Biotium, Inc. Hayward, CA). Three markers (each 1 µl) with different labels were subsequently pooled

and denatured with 7 µl Hi-Di Formamid (Applied Biosystems). Fragment analysis was subsequently performed at the Sequencing Facility at Yale University using LIZ-500 as size standard.

Genotypes were determined from chromatograms manually using the microsatellite plug-in in Geneious. Micro-Checker (Van Oosterhout et al. 2004) was used to test our data for genotyping errors due to stutter and null alleles. After this quality control step the data was imported into Create 1.37 (Coombs et al. 2008) to generate input files for downstream population genetic analyses.

We tested our data for departures from Hardy-Weinberg equilibrium and linkage disequilibrium using Genepop on the Web v. 4.2 (Raymond & Rousset 1995). Statistical significance of the tests were evaluated after Bonferroni correction for multiple comparisons. The number of alleles, observed and expected heterozygosities, and allelic richness were calculated with FSTAT v. 2.9.3.2 (Goudet 1995). Estimates of neutral divergence were calculated as F_{ST} in FSTAT using the Weir & Cockerham (1984) method. In addition, we calculated a second estimate of genetic differentiation Jost's D (D_{est}) using the online application SMOGD v. 1.2.5 (Crawford 2010). For both estimates of population divergence we tested for isolation by distance using the IbD web server v. 3.25 (Jensen et al. 2005). Estimates of effective population sizes were calculated based on gametic disequilibrium between alleles from different neutral loci as implemented in the LD option in N_eEstimator (Ovenden et al. 2007).

We estimated the number of genetic clusters supported by our microsatellite data using the Bayesian approach implemented in Structure v. 2.3.4 (Pritchard et al. 2000, Falush et al. 2003). We used the bulk run function and ran analyses for K=1 to K=10 with

10 replicates each. Each of the runs was performed for 1 million iterations discarding a burn-in of 100,000 samples. The resulting output was imported into Structure Harvester v. 0.6.93 (Earl & von Holdt 2012) which was used to estimate the most likely K for the data using the ΔK method proposed by Evanno et al. (2005). The program also provides the input files for the program CLUMPP v. 1.1.2 (Jakobsson & Rosenberg 2007) which was used to align the cluster assignments across the 10 replicate runs for each K. Finally the data for K = 2, 3, 4, and 10 were visualized using Distruct v. 1.1 (Rosenberg 2004).

Migration rates were estimated in two different ways: First, we used F_{ST} to estimate migration rates using the relationship $M = (1/F_{ST}-1)/4$. In addition we used the program migrate-n (Beerli & Palczewski 2010) which implements a Bayesian approach to calculate migration rates. We used the Stepwise Mutation Model (SSM, Ohta & Kimura 1973) and ran the simulation for 1 million iterations and discarded a burn-in of 10,000 steps. Default settings were used for the remaining parameters.

Comparative Analyses of Population Divergence

Two methods are commonly used to calculate phenotypic divergence, Q_{ST} and P_{ST}. Q_{ST} is defined as the among-population proportion of the additive genetic variance of a quantitative trait as such it is not equivalent to the phenotypic variance among natural populations (the phenotypic variance will include non-additive variance in addition to the additive genetic variance) (Spitze 1993, Sæther et al. 2007). Hence, Q_{ST} estimates require that no environmental variability influences the phenotypes, which is easiest achieved using a common garden experiment. This, however, is not possible for many organisms, especially if a large number of natural populations are studied. The phenotypic differentiation, P_{ST}, is analogous to Q_{ST}, but is less sensitive to the influence of

environmental and non-additive genetic effects and therefore is readily applicable to natural populations (Merilä & Crnokrak 2001, Leinonen et al. 2006, Magalhaes et al. 2009). We calculated phenotypic variance, P_{ST}, as follows:

$$P_{ST} = \frac{\sigma^2{}_{GB}}{\sigma^2{}_{GB} + 2\,(h^2\sigma^2{}_{GW})},$$

We assumed $h^2 = 0.5$ for all traits. This assumes that half of the variation between populations is due to environmental effects and non-additive genetic variance. This has been suggested to be a reasonable estimate for a variety of morphological phenotypes (Leinonen et al. 2006, Magalhaes et al. 2008). In addition, a recent comparison of field and laboratory reared *M. zebra* to investigate the genetics of body shape and melanophore number suggested that both characters had a strong genetic component (Husemann et al. in review, Appendix 4.1). Therefore, $h^2 = 0.5$ can be considered a conservative estimate. For P_{ST} calculations data for all traits were adjusted for sex and size differences (besides standard length) by using the residuals from a preparatory analysis of variance. P_{ST} calculations were performed using custom R-scripts.

We calculated neutral genetic divergence as F_{ST} from our microsatellite data and Φ_{ST} from the sequence data. We use this classic F_{ST} estimate of genetic divergence and not the adjusted divergence estimate (D_{EST}) as F_{ST} is calculated in a similar fashion as P_{ST} (Edelaar et al. 2011). In addition we compared P_{ST} to Φ_{ST} as microsatellites mutation rates have been suggested to be too high for this comparison (Edelaar et al. 2011).

Pearson correlations were performed to investigate the relationships between genetic and phenotypic divergence matrices. We further investigated correlations between the phenotypic traits. All population comparisons were plotted to visualize

individual differences of genetic divergence estimates (Φ_{ST}, F_{ST}) and phenotypic divergent estimates (P_{ST}).

Results

Phenotypic Data

Most populations were significantly differentiated in body shape (Fig. 4.3). PCA yielded 20 axes explaining more than 95 % of the variance. Four significant population divergence vector axes were detected. MANCOVA revealed significant effects of standard length, sex and population identity on body shape (Table 2). The population identity had the highest effect with $\eta^2_p = 0.255$, only slightly larger than the effect of standard length ($\eta^2_p = 0.253$). Sex had less effect ($\eta^2_p = 0.148$). The interaction term of sex and population identity was not significant ($p = 0.052$).

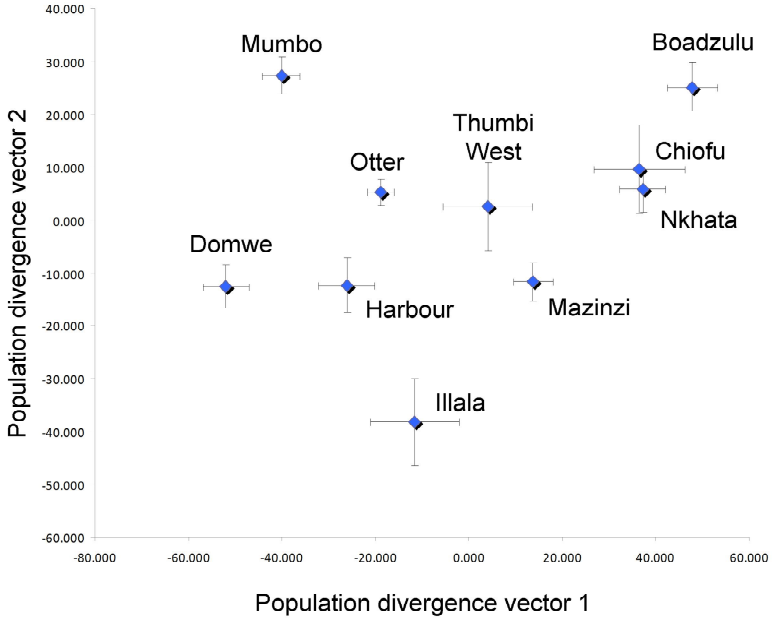

Figure 4.3. Body shape distribution of the 10 studied populations of *M. zebra*; displayed are the population divergence scores 1 and 2.

Table 4.2. Results of the multivariate analysis of covariance (MANCOVA) on the population divergence scores. *F*-ratios were approximated using Wilks' lambda, effect sizes were estimated using Partial Eta squared (η^2_p) and analyses of variance (ANOVA) for standard length and melanophore data. Melanophore data was only collected for males, hence we did not test for a sex effect.

Effect	*F*	Hypothesis d.f.	Error d.f.	p-value	η^2_p
MANCOVA body shape					
Intercept	4.695	20	253.0	<0.001	0.271
Standard length	4.276	20	253.0	<0.001	0.253
Sex	2.189	20	253.0	0.003	0.148
Population	4.438	180	2105.4	<0.001	0.255
Population*Sex	1.198	160	1903.6	0.052	0.086
ANOVA standard length	*F*	d.f.	SS	p-value	
Sex	58.020	1	2397	<0.001	
Population	26.323	9	9787	<0.001	
Sex* Population	1.629	8	538	0.116	
ANOVA melanophores					
Standard length	3.997	1	1895	0.0476	
Population	13.835	9	59021	<0.001	
Standard length*Population	0.286	9	1221	0.9775	

The standard deviations of the standard length data overlapped across populations (Fig. 4.4b). Means differed between 91.9 mm for the Illala Gap population and 71.4 mm for the Nkhata Bay population. ANOVA yielded significant population effects ($p < 0.001$, Table 4.2). Further a significant effect of sex was detected for the standard length data ($p < 0.001$). Tukey's HSD found significant differentiation for 21 of the 45 pairwise comparisons after adjusting for sex ($\alpha = 0.05$).

Most populations were fairly similar in their melanophore counts. Differences were mostly found between the populations from Domwe Island and Otter Point and the remaining populations (Fig. 4.4a). Means varied between 57.1 (± 13.1 SD) for Otter Point to 119.4 (± 14.3 SD) for Nkhata Bay (Fig. 4.4a).

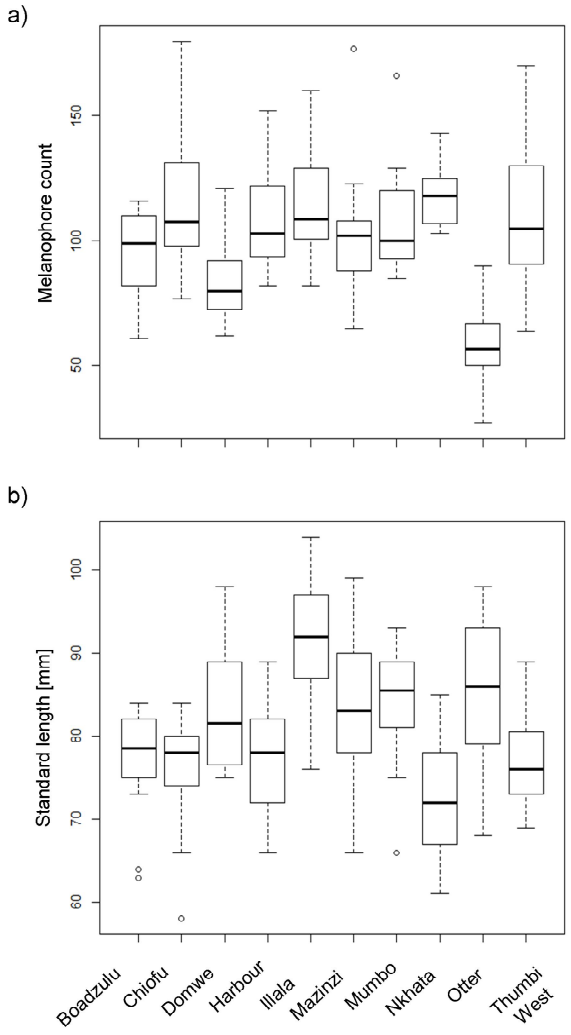

Figure 4.4. Box plots showing a) the distribution of melanophore counts and b) the distribution of standard lengths across the studied populations.

The populations from Otter Point had significantly lower counts than all other populations, whereas the Domwe Island population had lower counts than most populations except than the Boadzulu Island population (Tukey's HSD, $p < 0.05$). Further, the populations from Nkhata Bay and Boadzulu Island were significantly different in their melanophore counts with the Nkhata Bay population having less melanophores than the Boadzulu Island population.

Mitochondrial DNA

A total of 332 sequences with a length of 646 bp were obtained for the ten populations (21-47 sequences per population, Table 4.3).

Table 4.3. Summary statistics for the mtDNA data for all populations.

Pop	# seq	Polymorph sites (S)	# haps	Hd	K
Domwe	27	7	6	0.63818	1.86325
Thumbi	23	4	6	0.77075	1.21739
Illala	37	9	8	0.68619	1.20120
Otter	23	13	9	0.84585	1.83399
Chiofu	24	18	9	0.70652	3.78623
Harbour	47	12	12	0.76688	4.35985
Boadzulu	21	5	6	0.69524	0.93333
Mazinzi	45	7	6	0.54747	1.61818
Mumbo	28	2	3	0.31481	0.37566
Nkhata	27	8	8	0.73789	1.92023
Total	302	46	57	0.92676	4.25929

A total of 57 haplotypes were recovered, resulting in a total haplotype diversity of Hd = 0.927 and a nucleotide diversity of Pi = 0.006. The most diverse populations were Otter Point and Harbour Island with haplotype diversities of Hd = 0.846 and Hd = 0.767, respectively. The least diverse populations were found at Mumbo Island and Mazinzi Reef with Hd = 0.315 and Hd = 0.547, respectively. Φ_{st} were mostly significant with the exception of the value between Otter Point and Illala Gap. Most estimates were relatively high and ranged between 0.0175 between Otter Point and the Illala Gap and 0.854 between Mumbo Island and Thumbi West Island (Table 3). No relationship between linear geographic distance and genetic distance was detected (Fig. 4.5a, Z = 1764.07, r = -0.007, one sided p = 0.51).

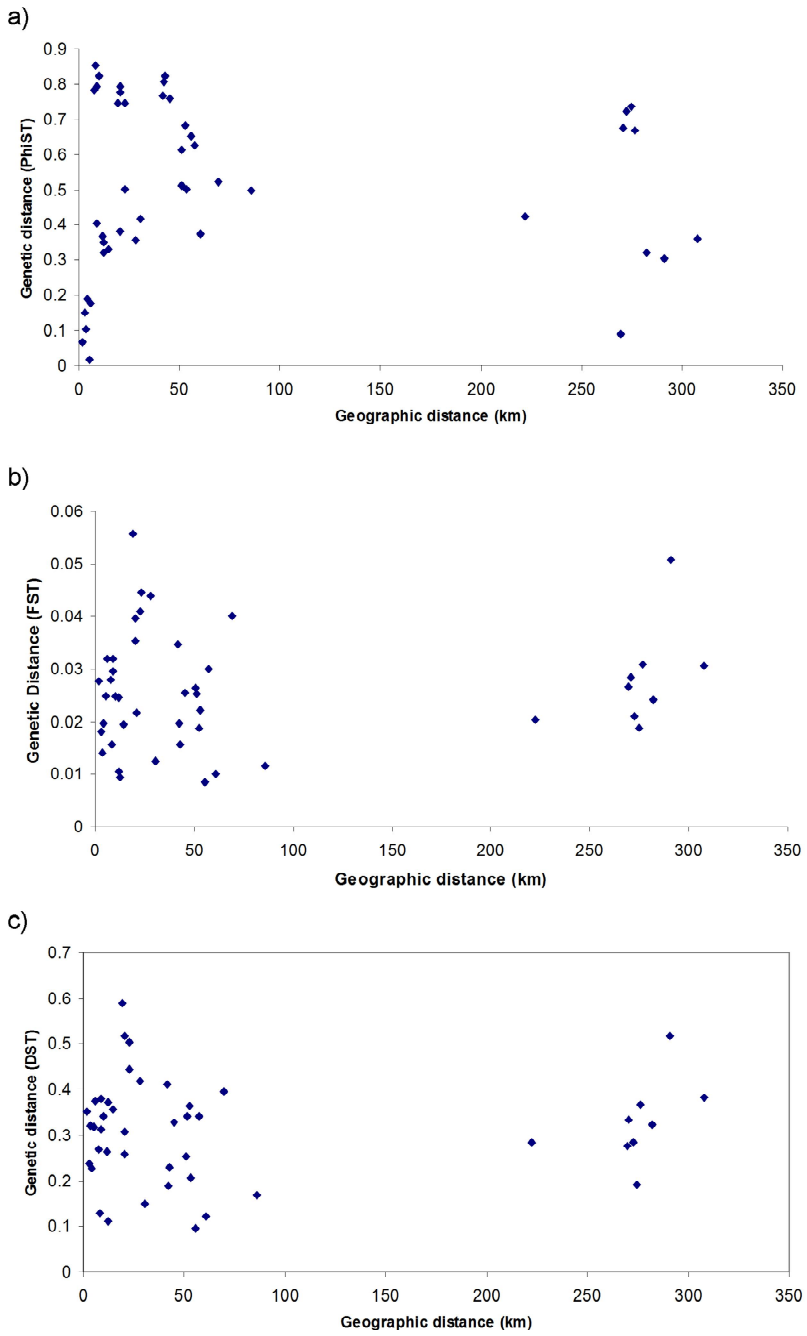

Figure 4.5. Isolation by distance a) using Φ_{ST} from mitochondrial data, b) using F_{ST} calculated from microsatellite data, c) using D_{EST} from microsatellite data generated with IbD web service v. 3.23; none of the analyses showed a significant isolation by distance pattern; significance was tested using 1000 permutations.

Bayesian skyline analyses suggested that most populations expanded within the last 20,000 years (Fig. 4.6). Many populations could not be traced back in time further than 10,000 years. Extant female population sizes differed by a factor of three with populations from Nkhata Bay and Harbour Island having the largest female population sizes and populations from Mumbo Island, Illala Gap, Mazinzi Reef and Domwe Island being the smallest.

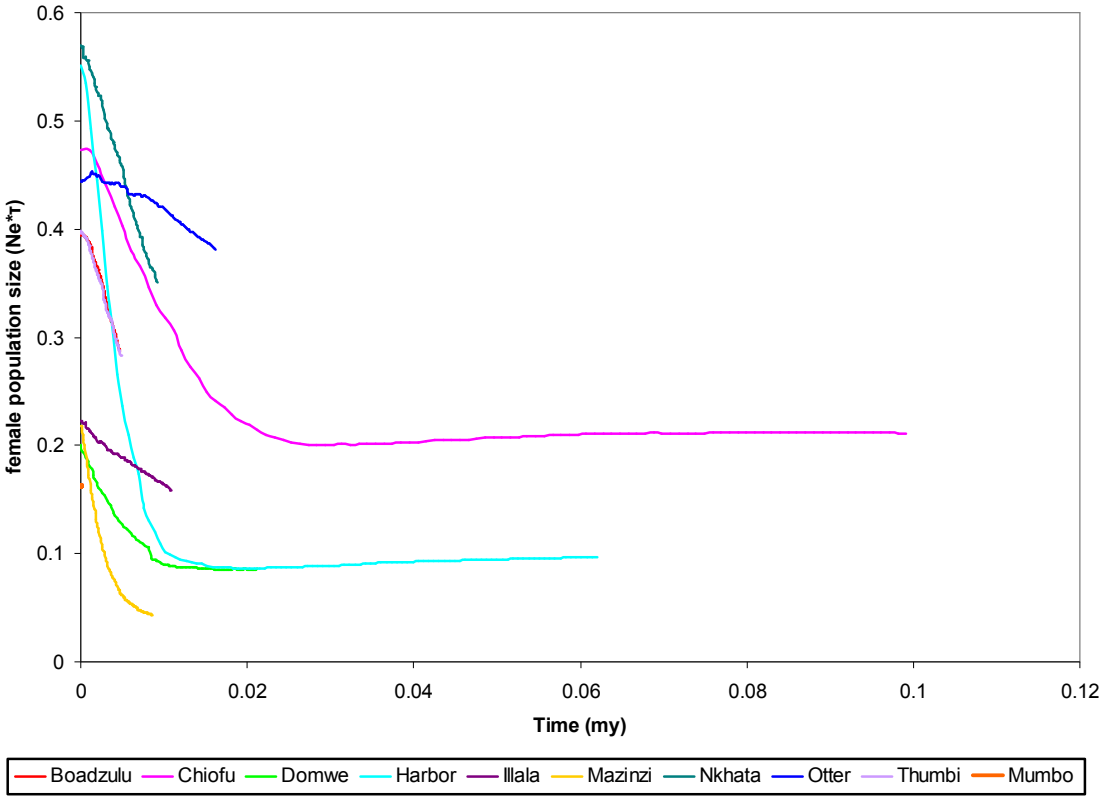

Figure 4.6. Bayesian Skyline plots to visualize the demographic history of the studied populations.

Microsatellites

Of the 14 loci originally chosen for the study 3 did not amplify well for many individuals and were discarded. An additional locus was removed from the analyses

because an excess of homozygotes was detected for almost all populations (UNH362). Therefore, only 10 microsatellite loci were used for subsequent analyses (Appendix 3.1). An average of 37 (range: 27 - 60) individuals were genotyped for each of the 10 populations of *M. zebra* (Table 4). Micro-Checker revealed no evidence of distortion by stutter or allele drop out for any locus. Slight excess of homozygotes was detected for single loci at some populations (Table 4). Deviations from HWE after Bonferroni correction were found for the loci UNH2065 and UNH231 for the Boadzulu Island and were likely due to the presence of null alleles as detected by Micro-Checker. Global analyses of linkage disequilibrium suggested that no significant linkage existed for any of the loci.

Table 4.4. Summary statistics for microsatellite data.

location	# alleles	Average # alleles per locus	Ho	He	Average allelic richness	Deviation HWE	Possible null alleles	Ne (N_eEstimator, LD)
Domwe	200	20.0 (SD 3.9)	0.886	0.931	18.219	-	-	165.9 (113.0 – 303.1)
Thumbi	231	23.1 (SD 5.4)	0.889	0.944	21.351	-	UNH2152	169.1 (123.9 – 262.1)
Illala	206	20.6 (SD 3.4)	0.923	0.937	18.362	-	-	246.3 (159.2 – 525.4)
Otter	197	19.7 (SD 5.8)	0.856	0.930	18.367	-	-	313.2 (173.6 – 1414.9)
Chiofu	220	22.0 (SD 5.1)	0.913	0.941	19.884	-	-	173.1 (125.8 – 272.7)
Harbor	250	25.0 (SD 5.1)	0.868	0.952	21.852	-	UNH2037 UNH2204	355.3 (218.3 – 916.2)
Boadzulu	221	22.1 (SD 4.0)	0.874	0.939	20.475	UNH2065 UNH231	UNH2204 UNH2065 UNH231	208.5 (141.3-387.6)
Mazinzi	165	16.5 (SD 4.3)	0.863	0.900	14.733	-	-	80.6 (64.5 – 106.2)
Mumbo	228	22.8 (SD 4.6)	0.889	0.936	20.691	-	-	235.8 (153.6 – 491.5)
Nkhata	228	22.8 (SD 6.3)	0.886	0.934	20.557	-	UNH2169	203.2 (141.1 – 355.6)

The number of alleles across all loci ranged between 165 for Mazinzi Reef and 250 for Harbour Island, which was the genetically most diverse population with an allelic richness of 21.85. Mazinzi Reef had the lowest allelic richness with 14.73 (Table 4). All populations were significantly differentiated from each other, indicated by significant F_{ST} values and high estimates of D_{EST} (Table 5b). Both estimates were strongly correlated (R^2 = 0.86).

Table 4.5a: Φ_{ST} for mitochondrial data calculated with Arlequin under the Tamura-Nei model. bold font indicates significance at $\alpha = 0.05$.

Pop.	Domwe	Thumbi	Illala	Otter	Chiofu	Harbor	Boadzulu	Mazinzi	Mumbo	Nkhata
Domwe	0									
Thumbi	**0.191**	0								
Illala	**0.149**	**0.104**	0							
Otter	**0.176**	**0.067**	0.018	0						
Chiofu	**0.613**	**0.653**	**0.683**	**0.626**	0					
Harbor	**0.332**	**0.353**	**0.367**	**0.320**	**0.375**	0				
Boadzulu	**0.760**	**0.824**	**0.806**	**0.766**	**0.500**	**0.418**	0			
Mazinzi	**0.745**	**0.791**	**0.777**	**0.745**	**0.523**	**0.404**	**0.502**	0		
Mumbo	**0.783**	**0.854**	**0.823**	**0.793**	**0.502**	**0.382**	**0.511**	**0.359**	0	
Nkhata	**0.674**	**0.735**	**0.723**	**0.670**	**0.425**	**0.321**	**0.361**	**0.306**	**0.090**	0

Table 4.5b: F_{ST} for microsatellite data (above diagonal). All F_{ST} estimates were significantly greater than zero ($\alpha = 0.05$). D_{EST} estimates are the below diagonal. D_{EST} significance not assessed.

Pop.	Domwe	Thumbi	Illala	Otter	Chiofu	Harbor	Boadzulu	Mazinzi	Mumbo	Nkhata
Domwe	0	0.028	0.028	0.039	0.033	0.028	0.033	0.052	0.030	0.035
Thumbi	0.227	0	0.023	0.029	0.008	0.012	0.016	0.036	0.017	0.019
Illala	0.239	0.319	0	0.029	0.027	0.020	0.026	0.049	0.031	0.027
Otter	0.374	0.351	0.318	0	0.031	0.027	0.035	0.055	0.029	0.033
Chiofu	0.341	0.096	0.364	0.340	0	0.010	0.012	0.040	0.022	0.021
Harbor	0.358	0.110	0.263	0.371	0.122	0	0.012	0.034	0.023	0.024
Boadzulu	0.329	0.230	0.189	0.412	0.167	0.151	0	0.040	0.027	0.030
Mazinzi	0.503	0.308	0.517	0.589	0.396	0.379	0.444	0	0.044	0.051
Mumbo	0.269	0.130	0.342	0.313	0.206	0.257	0.253	0.419	0	0.027
Nkhata	0.332	0.192	0.285	0.367	0.283	0.322	0.382	0.518	0.276	0

The strongest differentiations were found between Mazinzi Reef and most other populations. This was also confirmed by our analyses of genetic clusters in Structure. The analysis yielded strong support for K = 2, where one cluster was formed by the Mazinzi Reef population and a second cluster by all other populations (Fig. 4.7, 4.8). No relationship between linear geographic distance and genetic distance was detected for either classic F_{ST} (Fig. 4.5b, Z = 101.15, r = 0.007, one sided p = 0.44) or D_{EST} (Fig. 4.5c, Z = 1108.93, r = 0.059, one sided p = 0.34).

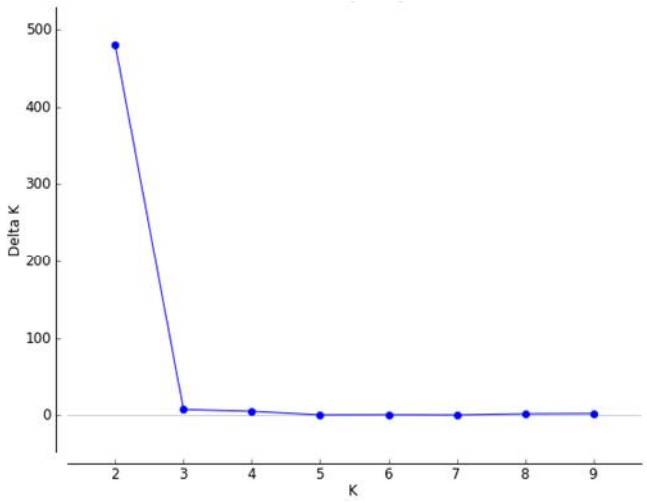

Figure 4.7. ΔK distinguishing the most likely number of clusters for Structure analysis of the ten populations of *M. zebra* calculated with Structure Harvester.

This is also reflected in the estimated migration rates. Migration was lowest between Mazinzi and any other location with 4-7 migrants per generation calculated from F_{ST} estimates (Table 6). The highest migration estimates were found between Chiofu and Thumbi (21 migrants per generation) and Chiofu and Thumbi West Island (30 migrants per generation). The estimates from migrate-n were much higher but showed similar patterns (Table 6).

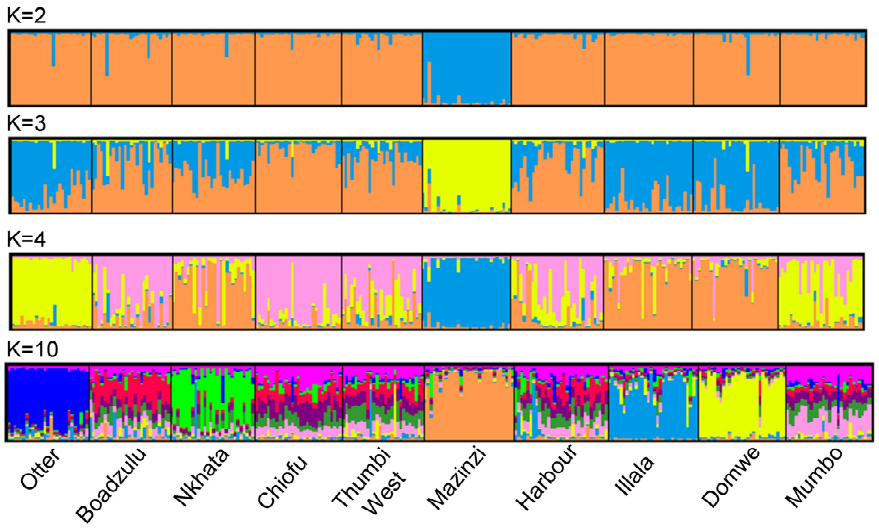

Figure 4.8. STRUCTURE results for K = 2, K = 3, K = 4 and K = 10. Analysis of ΔK yielded the highest support for K = 2.

Estimates of effective population sizes (N_e) ranged between 81 Individuals for Mazinzi Reef to 355 Individuals for Harbour Island; N_e for most populations ranged between 150 and 250 individuals (Table 4).

Table 4.6. Migration rates between the subpopulations of *Maylandia zebra* calculated from FST estimates (below diagonal) and two way migration rates estimated with migrate-n (above diagonal).

Pop.	Boadzulu	Chiofu	Domwe	Harbour	Illala	Mazinzi	Mumbo	Nkhata	Otter	Thumbi
Boadzulu	0.000	53.51	17.30	25.452	26.61	17.623	20.84	26.08	10.9	68.94
		64.32	3.850	68.777	28.81	97.228	6.053	27.98	38.95	78.70
Chiofu	20.55	0.000	38.22	27.374	34.25	12.696	92.756	19.95	90.71	26.94
			33.77	30.766	8.60	7.086	10.801	78.64	25.66	29.29
Domwe	7.276	7.235	0.000	22.087	27.88	65.518	11.093	24.55	15.52	57.31
				59.000	38.20	10.764	14.359	19.07	27.32	59.79
Harbour	20.18	24.31	8.650	0.000	13.66	22.854	13.855	17.55	15.95	46.40
					87.52	31.671	34.301	15.83	77.61	88.28
Illala	9.474	9.096	8.669	12.300	0.000	52.485	33.407	13.92	135.86	23.65
						7.207	42.868	44.90	35.06	15.11
Mazinzi	5.941	5.941	4.599	7.032	4.838	0.000	20.567	8.889	24.10	28.5
							10.242	19.59	24.67	5.356
Mumbo	9.099	10.931	8.139	10.827	7.822	5.468	0.000	101.03	16.34	39.63
								31.26	4.56	23.34
Nkhata	8.014	11.695	6.967	9.913	8.938	4.633	9.145	0.000	28.12	21.33
									14.23	45.15
Otter	6.928	7.807	6.123	8.931	8.245	4.322	8.274	7.442	0.000	24.67
										33.23
Thumbi	15.704	29.80	8.565	20.259	10.64	6.671	14.902	12.89	8.338	0.000

Comparative Analyses

The estimates of phenotypic differentiation (P_{ST}) were high for most population comparisons for all three traits (Table 4.7a, b, Fig. 4.9, 4.10). All P_{ST} estimates for body shape were significant and large; only a single estimate was below 0.9. P_{ST} for most comparisons of body length were above 0.8. For melanophore counts a larger number of lower estimates (<0.8) were found. However, for all phenotypic traits the medians of P_{ST} across all 45 pairwise population comparisons were larger than F_{ST} (Fig. 4.9, Tukey's HSD, $p < 0.001$).

Table 4.7a. P_{ST} values for melanophore data (below diagonal); significant estimates are printed in bold ($\alpha = 0.05$).

Pop.	Domwe	Thumbi	Illala	Otter	Chiofu	Harbor	Boadzulu	Mazinzi	Mumbo	Nkhata
Domwe	0									
Thumbi	**0.985**	0								
Illala	**0.990**	**0.997**	0							
Otter	**0.990**	0.313	**0.998**	0						
Chiofu	**0.990**	0.610	0.344	**0.998**	0					
Harbor	**0.986**	0.007	0.381	**0.997**	0.652	0				
Boadzulu	0.898	**0.968**	**0.981**	**0.994**	**0.981**	**0.971**	0			
Mazinzi	0.963	0.946	0.969	**0.997**	**0.972**	0.946	0.806	0		
Mumbo	**0.976**	0.086	0.446	**0.995**	0.643	0.062	0.952	0.902	0	
Nkhata	**0.992**	0.848	0.800	**0.998**	0.587	0.874	**0.988**	**0.981**	0.861	0

Correlations between Φ_{ST} and P_{ST} were low for all traits ($R^2 < 0.02$, Table 4.8). When regressing F_{ST} against P_{ST} a small but significant correlation was found for melanophore counts ($R^2 = 0.187$, $p = 0.002$) and body shape ($R^2 = 0.115$, $p = 0.013$), yet the correlation was non-significant for standard length ($R^2 = -0.021$, $p = 0.756$). No correlation was found between P_{ST} for standard length and body shape ($R^2 = 0.020$, $p = 0.174$). However, a slight, but significant correlation was found between melanophore counts and body shape ($R^2 = 0.0378$, $p = 0.038$).

Table 4.7b. P_{ST} values for standard length data (below diagonal) and body shape based on 20 PC axes (above diagonal); significant estimates are printed in bold ($\alpha = 0.05$).

Pop.	Domwe	Thumbi	Illala	Otter	Chiofu	Harbor	Boadzulu	Mazinzi	Mumbo	Nkhata
Domwe	0	**0.933**	**0.948**	**0.955**	**0.967**	**0.927**	**0.967**	**0.959**	**0.951**	**0.956**
Thumbi	**0.978**	0	**0.933**	**0.935**	**0.935**	**0.924**	**0.932**	**0.955**	**0.941**	**0.921**
Illala	**0.976**	**0.996**	0	**0.952**	**0.944**	**0.894**	**0.949**	**0.941**	**0.950**	**0.921**
Otter	0.867	**0.991**	**0.957**	0	**0.954**	**0.923**	**0.959**	**0.951**	**0.946**	**0.930**
Chiofu	**0.979**	0.222	**0.995**	**0.991**	0	**0.954**	**0.936**	**0.958**	**0.969**	**0.940**
Harbor	**0.958**	0.777	**0.993**	**0.984**	0.840	0	**0.950**	**0.928**	**0.931**	**0.933**
Boadzulu	**0.960**	0.821	**0.994**	**0.985**	0.870	0.029	0	**0.949**	**0.956**	**0.931**
Mazinzi	0.134	**0.979**	**0.985**	**0.928**	**0.981**	**0.955**	**0.956**	0	**0.966**	**0.939**
Mumbo	0.126	**0.976**	**0.980**	0.905	**0.978**	**0.953**	**0.955**	0.000	0	**0.953**
Nkhata	**0.989**	**0.913**	**0.997**	**0.995**	0.866	**0.962**	**0.968**	**0.990**	**0.989**	0

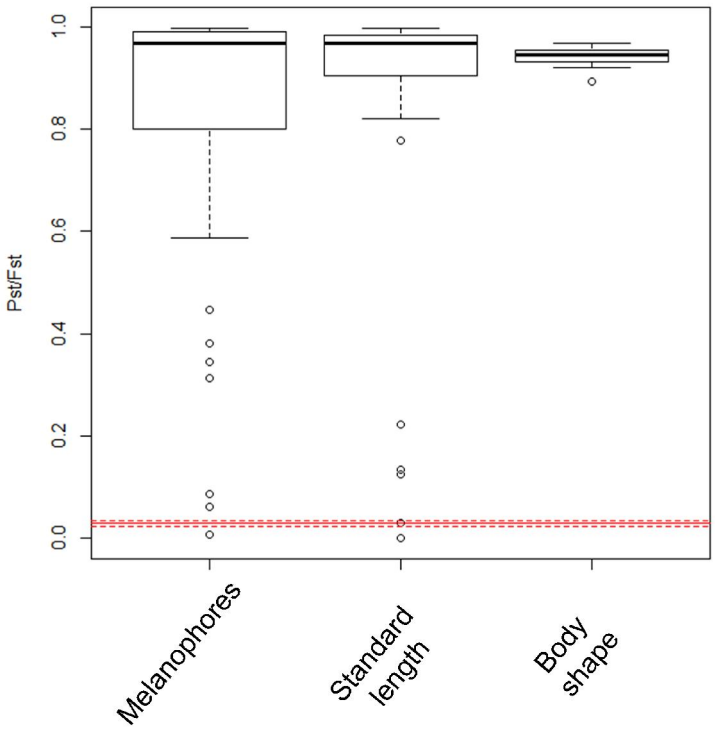

Figure 4.9. Comparison of neutral genetic divergence (F_{ST}, red lines, media – solid line, 95% CI broken lines) and phenotypic divergence (P_{ST}, boxes) for melanophore counts, standard length and body shape across all population comparisons.

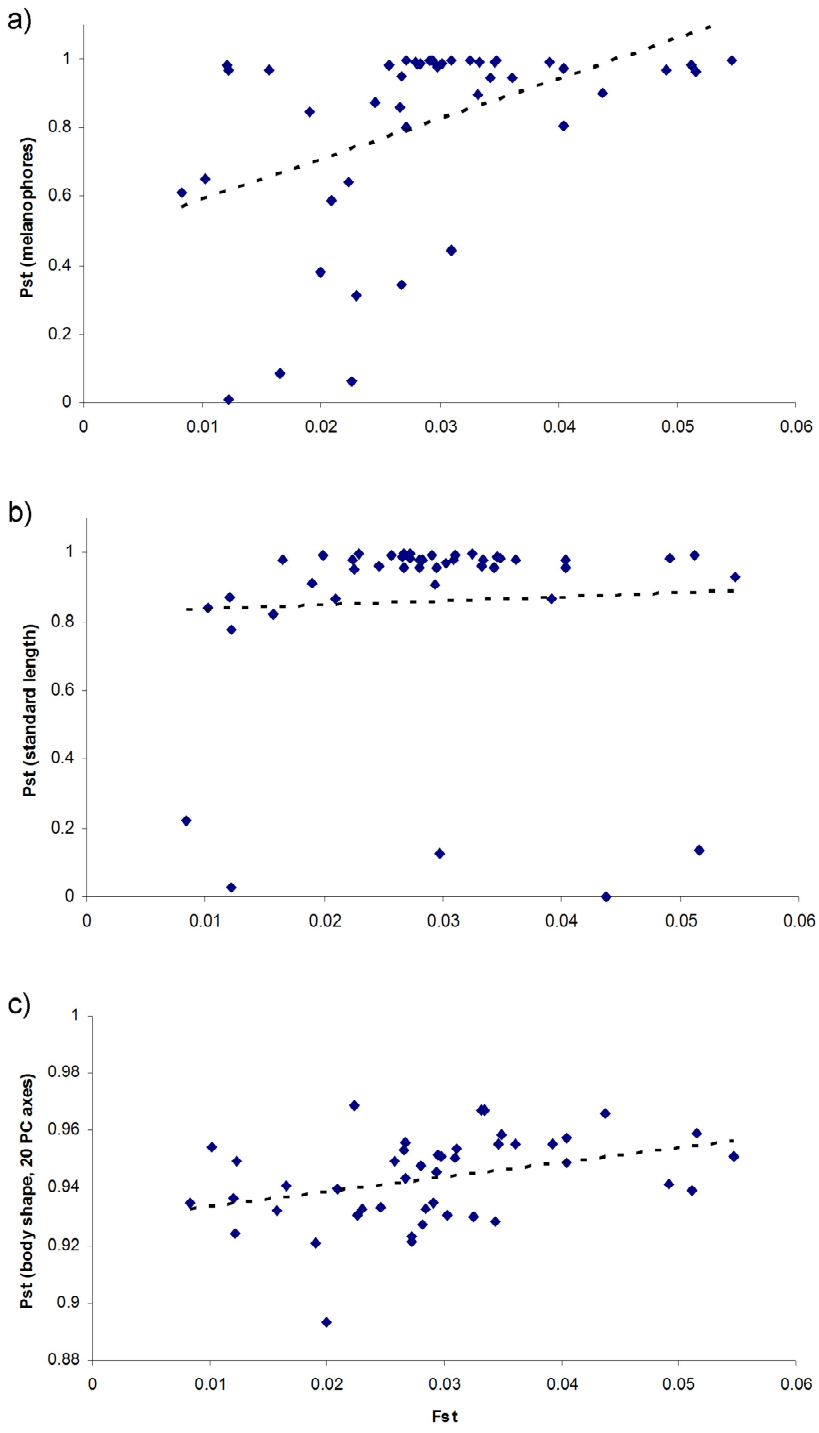

Figure 4.10. Comparison of neutral genetic divergence (F_{ST}) and the phenotypic divergence (P_{ST}) for a) melanophore counts, b) standard length and c) body shape in all individuals; a trend line is given for each comparison.

Table 4.8. Correlations (R^2) between pairwise P_{ST} at single phenotypic traits and pairwise F_{ST} from ten putative neutral microsatellite loci or pairwise Φ_{ST} from mitochondrial D-Loop sequences.

Trait	R^2_{usat}	p-value	R^2_{mtDNA}	p-value
Melanophores	**0.1873**	**0.001752**	-0.02316	0.9502
Standard length	-0.02094	0.7563	-0.01337	0.5207
Body shape	**0.1146**	**0.01314**	0.01318	0.2144

Discussion

Population Structure, Demographic History and Population Sizes

Lake Malawi's rock-dwelling cichlids are known to exhibit significant genetic differentiation even over small geographic distances (Danley et al. 2000, Smith & Kornfield 2002, Streelman et al. 2007). This is consistent with our results; both the mitochondrial and microsatellite data support significant differentiation among all populations (Table 5a, b). However, the differentiation does not follow a geographic pattern and some of the geographically most proximate populations are the most diverged (Table 4.5, Fig. 4.5). For the mitochondrial data this might be partially a result of incomplete lineage sorting. For the microsatellite data the high genetic diversity and the relatively low sample sizes might contribute to this pattern as genetic diversity might be undersampled. Alternatively, some of the geographically close but strongly diverged populations might represent cryptic species; such patterns of cryptic diversity are common in the rock-dwelling cichlids and new species with subtle morphological differences are frequently discovered (Stauffer et al. 1997, 2013, Ciccotto et al. 2011).

Our structure analysis on microsatellite data found two different genetic clusters (Fig. 4.8): the first cluster consists of the population from Mazinzi Reef, whereas all other populations combined, represent the second cluster. Several reasons might explain the

lack of isolation by distance and the genetic cluster pattern assuming the neutral evolution of the markers. On the one hand, the Mazinzi Reef population is strongly isolated and has a very low population size. The reef also completely desiccated in the last 1,500 years (Owen et al. 1990), making the colonization of the reef extremely recent and elevating the likelihood of founder effects playing a role in this population's divergence. Furthermore, Mazinzi Reef is relatively small (~10,000 m^2) and far from the nearest rocky habitat both of which contribute to this population's isolation. Therefore, strong drift my account for the uniqueness of the population.

On the other hand, this reef is relatively untouched by humans compared to many of the other sample sites. It is possible that the divergences between many other locations were naturally higher but this divergence has been erased by the introduction and hybridization with non-native species and individuals of other conspecific populations. Such introductions have been documented at Nkhata Bay, Thumbi West Island (Genner et al. 2006), and more recently, the immigration of species beyond their historic range has been observed at Harbour Island (P.D.D. personal observation). Furthermore, hybridization between native and translocated species is considered common (Streelman et al. 2004, Zidana et al. 2009). Several of our other sampling locations in the southern part of the lake are close to Thumbi West Island and may have received migrants from species introduced to Thumbi West Island. The accidental release of specimens at Chiofu Bay is possible as well, as a fish exporter is located at this location. Therefore, natural patterns of divergence might be distorted.

However, as previously stated, all populations are significantly diverged from each other. This might be at least partially the result of stochastic events in small

populations. The N_e for most populations are relatively small (Table 4) and many populations strongly expanded in recent times (Fig. 4.6). The small populations and their rapid expansion might have resulted in divergence driven by founder effects and drift. However, the N_e estimates derived from methods using single temporal samples are generally smaller than estimates from multiple time points (Barker 2011) and estimates from multiple time points for some of the studied populations are much higher (Husemann et al. in prep). Therefore, drift might be only of limited importance for widespread species in Lake Malawi (Danley et al. 2000, Husemann et al. in prep.).

Selection and Phenotypic Evolution

In the past, the traits we investigate here have been suggested to be under selection (Streelman et al. 2007, Salzburger 2009, Husemann et al. in review); yet, these hypotheses were mostly based on observational data rather than empirical tests.

We used the comparison of phenotypic differentiation (P_{ST}) and neutral genetic divergence (F_{ST}, Φ_{ST}) to test if phenotypes are under selection. However, the results of such tests need to be interpreted with caution, as the $F_{ST} - P_{ST}$ method has many caveats. P_{ST} might be downwardly biased if sample sizes are relatively low (O'Hara & Merilä 2005). Further, P_{ST} depends on the estimated heritability. A commonly used value assumed for morphological traits is $h^2 = 0.5$ (Leinonen et al. 2006, Magalhaes et al. 2009). Though heritability was not calculated for the measured phenotypes, both body shape and melanophore count data from 12^{th} generation laboratory reared individuals and wild caught specimens were largely consistent indicating high heritability (Appendix 4.1). Therefore we consider $h^2 = 0.5$ a conservative estimate. The F_{ST} estimates may also be problematic: in species with high rates of gene flow rates, F_{ST} might be overestimated

(Waples 1998). Our analyses indicated that migration rates between the studied populations might be high (Table 6). In addition, high mutation rates of microsatellite loci might be problematic as that reduces F_{ST} estimates (Edelaar et al. 2011). For this reason we compared our P_{ST} estimates to both, F_{ST} from microsatellites and Φ_{ST} from mitochondrial sequences and strong conclusions were only made if patterns hold up for both estimates of neutral divergence.

The population differentiation in standard length and body shape exceeds neutral expectations derived from microsatellites by an order of magnitude. When compared to Φ_{ST}, P_{ST} still exceeds neutral expectations, yet not as clearly. This suggests that strong local selection drives phenotypic evolution in these two traits as it has been suggested in the past (Streelman et al. 2007, Salzburger 2009). This is not surprising as each location represents a slightly different selective environment. While the most physico-chemical parameters are fairly similar across the lake, it has been shown that resources and community composition strongly differ between locations (Reinthal 1990, Abdallah & Barton 2003, Higgins et al. 2003, Ribbink et al. 1983, Parnell & Streelman 2011, Ding et al. in review). Even slight differences in the selective environment might lead to the divergence of populations in adaptive traits. The genetic diversity observed in most populations would indicate a high adaptability of the species (Loh et al. 2008). However, the influence of drift cannot be completely excluded as we observed a weak, but significant correlation between F_{ST} and P_{ST} for body shape. A potential influence of drift would be supported by the low estimates of N_e. Yet, as mentioned above, these are likely underestimates and meta-population sizes for *M. zebra* are likely much larger.

Populations are less strongly diverged in melanophore number and we observed a weak correlation of melanophore counts and F_{ST}. When compared to F_{ST} from microsatellites the divergence of melanophore counts still strongly exceeds neutral expectations. When compared to Φ_{ST} for some population pairs genetic divergence exceeds phenotypic divergence (Fig. 4.9). This may indicate either that selection on melanophore number is relatively weak, or that drift might play an important role for this character. Interestingly we observed a weak, but significant correlation between the divergence in body shape and melanophore counts. This might support the hypotheses that coloration and body shape might be under correlated selection (Husemann et al. in review).

Conclusions

Populations of *M. zebra* are strongly diverged in ecological characters and the divergence exceeds neutral expectations indicating strong local selection driving their divergence. Selection seems less strong on melanophore numbers, yet we observed a weak correlation between body shape and melanophore number providing further support for the potential of correlational selection on coloration traits and body shape.

.

CHAPTER FIVE

The Correlation of Reproductive and Ecological Traits May Facilitate the Coexistence of Congeners in Lake Malawi's Rock-Dwelling Cichlids

Introduction

Adaptive radiations are characterized by the rapid ecological divergence of lineages leading to a burst of speciation (Schluter, 2000). In vertebrates, it has been hypothesized that three stages of adaptive radiations can be distinguished (Streelman & Danley, 2003): the first stage is characterized by divergence in macrohabitat use, the second by trophic niche differentiation. While natural selection mainly drives divergence in the first two stages, sexual selection is thought to drive the diversification during the third stage where traits related to sexual communication diverge (Sturmbauer, 1998; Danley & Kocher, 2001; Salzburger, 2009). The differentiation of sexual signals alone among sympatric species during the third stage, however, leads to an ecological conundrum, as theory predicts ecological competition will lead to competitive exclusion (Koplin & Hoffmann, 1968; Armstrong & McGehee, 1980; Kaplan & Denno, 2007; Anderson 2008). Hence, the stable coexistence of species differing in sexual signaling should be accompanied by ecological trait divergence that mitigates interspecific competition among closely related species. Indeed, it has been shown that diversification in sexual signals can be associated with divergence in ecologically relevant traits in some adaptive radiations (Price, 1998; Boughman, 2001; Podos, 2001; Streelman & Danley, 2003; Servedio *et al*., 2011; Derryberry *et al*., 2012). Yet, such associations have

received little attention in some of the most diverse vertebrate radiations, such as parrotfish and cichlids (Streelman & Danley, 2003).

The haplochromine cichlids from East Africa are a model system for the study of adaptive radiation due to their young age and extreme morphological, ecological, and behavioral diversity (Seehausen, 2006; Salzburger, 2009; Sturmbauer *et al*., 2011). With more than 2000 species in the three East African Great Lakes (Tanganyika, Victoria and Malawi), cichlids represent the most species rich vertebrate radiations known (Genner *et al*., 2004; Koblmüller *et al*., 2008). Lake Malawi, with more than 700 endemic species of haplochromines, supports the most species rich species flock of East African cichlids (Kocher 2004; Danley *et al*., 2012a). The radiation of Lake Malawi's haplochromines appears to follow the three stage model of adaptive radiation: the initial split included differentiation into clades of rock- and sand dwellers, which appear to have hybridized soon after this split and given rise to the third major clade of Lake Malawi cichlids, the deep-benthic cichlids (Genner & Turner, 2012). This was followed by the divergence into genera occupying different trophic niches (Danley & Kocher, 2001). At the third stage, congeneric species diverged in sexual signaling traits, including chemical, acoustical, and visual characteristics (Allender *et al*., 2003; Blais *et al*., 2009; van Staaden & Smith, 2011; Danley *et al*., 2012b), and the female preferences for such traits (Plenderleith *et al*., 2005; Kidd *et al*., 2006). Congeneric species with different color phenotypes are often found in sympatry, and as many as five different species with different sexual signaling traits can coexist at a single location (Ribbink *et al*., 1983; Genner & Turner, 2005). This poses the question whether sympatric congeners differing primarily in sexual signaling traits have also diverged in ecologically relevant traits.

We studied the genus *Maylandia*, which represents the most species-rich genus of rock-dwelling cichlids in Lake Malawi. Species of this genus share similar body morphologies and differ predominantly in their male nuptial coloration (Stauffer *et al*., 1997, 2013; Allender *et al*., 2003). *Maylandia zebra,* the most widespread *Maylandia* species, is characterized by blue body coloration with distinctive black vertical bars along the body (Fig. 5.1, 5.2). In contrast, the majority of *Maylandia* species are endemic to a limited numbers of rocky outcrops within the lake where they are almost always sympatric with *M. zebra* (Ribbink *et al*., 1983). Several of these species are blue but lack the eponymous bars of *M. zebra* (Fig. 5.1, 5.2). A behavioral study suggested that, although the barred and non-barred species tend to occupy adjacent territories, their microhabitats may differ (Danley 2011).

The existence of multiple *M. zebra* populations with and without a sympatric non-barred congener offers the opportunity to test hypotheses about the role of fine scale ecological trait differentiation in mediating coexistence of closely related species. Previous research suggest that geographically distinct populations of *M. zebra* differ in body size and tropic morphology, indicating adaptation to local environmental conditions (Streelman *et al*., 2007); yet, that study did not test for the effects of sympatry with close relatives. Studies of sympatric species with alternative coloration phenotypes indicated that congeners exploit similar food sources (Martin & Genner, 2009), but can show slight differences in microhabitat use, territory size, and territory defense (Holzberg, 1978; Danley, 2011). Studies in other systems have shown that even such slight differences can lessen competitive pressure and help to allow for coexistence (Willis *et al*., 2005). Nonetheless, it remains unclear whether differential habitat use is reflected in phenotypic

trait divergence. Consequently, we studied variation in body shape among communities of *Maylandia* in which barred and non-barred species existed in sympatry and populations of *M. zebra* that lacked a sympatric non-barred congener.

In fish, body shape is known to respond to a variety of ecological sources of selection including predation (Langerhans, 2009; Ingram *et al*., 2012) and local abiotic conditions (Tobler & Carson, 2010; Tobler *et al*., 2011). In cichlids, body shape can be used to quantify even small differences in adaptive morphology (reviewed in Kerschbaumer & Sturmbauer, 2011). Thus we examined body shape to test whether divergence in sexual signaling is coupled with ecological trait divergence.

We sampled replicated communities of *Maylandia* in which the barred species either had or lacked a non-barred congener to test a series of hypotheses in regards to ecological differentiation between *Maylandia* species and among *M. zebra* populations. First, we used body shape as a proxy for ecological divergence in sympatric pairs of *Maylandia* species with barred and unbarred nuptial coloration to investigate if sympatric congeners differ in this ecological trait. Second, we tested the hypothesis that these differences are consistent across replicate species pairs found at different locations. Third, we used genetic markers and Bayes Factor analysis to investigate whether potentially consistent differences among replicated species pairs are the result of common ancestry or convergent evolution. Specifically, we tested whether taxa with similar color pattern are monophyletic. Fourth, we tested for morphological differences between populations of *M. zebra* that either were sympatric with a non-barred congener or lacked a non-barred congener. If ecological differentiation plays a role at the third stage of the radiation, we predict that the body shape of *M. zebra* populations sympatric with a non-barred species

differs from those that lack a non-barred congener, and that populations lacking the congener are morphologically more variable.

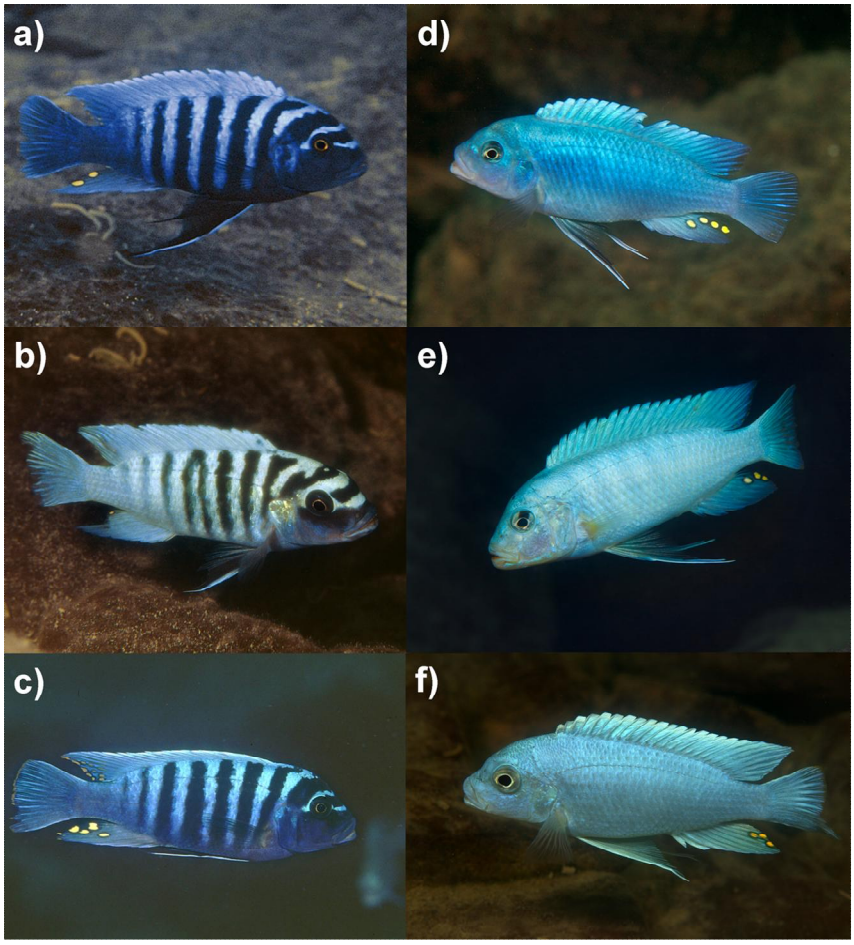

Figure 5.1. *Maylandia zebra* from a) Nkhata Bay, b) Mazinzi Reef, and c) Chiofu Bay and d) *M. callainos* from Luwino Reef (close to Nkhata Bay), e) *M. benetos* from Mazinzi Reef and f) *M. chrysomallos* from Gome (close to Chiofu Bay); pictures provided by Ad Konings).

Material and Methods

Specimen Collection

Our study focused on species within the genus *Maylandia*. The appropriate genus name for this group is contended and the species studied here have been classified as

belonging to the genus *Metriaclima* or, historically, *Pseudotropheus* (Regan, 1922; Stauffer *et al*., 1997, 2013). Here we follow the *Maylandia* designation (Meyer & Foerster, 1984; Condé, & Géry, 1999) as this name has been widely adopted in Genbank.

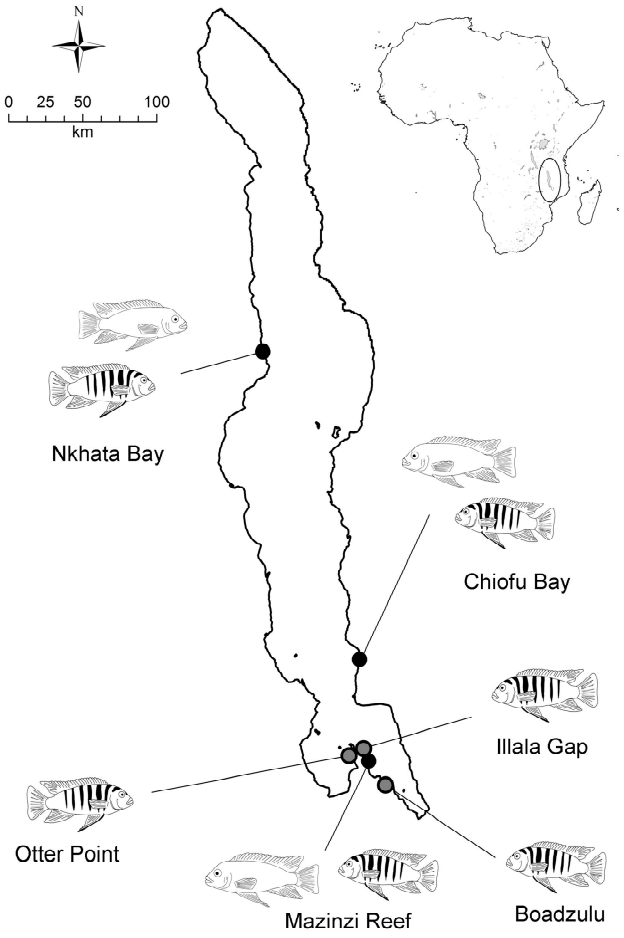

Figure 5.2. Map displaying the sampling sites for 'sympatric' species pairs of *Maylandia* (black dots) and 'allopatric' populations of *Maylandia zebra* (grey dots).

In 2010, we collected *M. zebra* at 6 locations (Nkhata Bay, Chiofu Bay, Mazinzi Reef, Boadzulu, Otter Point, Illala Gap, Table 5.1, Fig. 5.1, 5.2). At three of the locations we collected a sympatric non-barred species: *Maylandia benetos* (Mazinzi Reef), *Maylandia chrysomallos* (Chiofu Bay), and *Maylandia callainos* (Nkhata Bay). A total of

193 males were obtained. At some of the locations additional species of *Maylandia* can be found which are morphologically more divergent from *M. zebra* and were therefore not included in this study. In this manuscript, we refer to *M. zebra* populations that are sympatric with a blue congener as 'sympatric' populations and to populations of *M. zebra* that lack a sympatric blue congener as 'allopatric' populations. The distance between sampling locations varied between less than 2 km to more than 300 km (Fig. 5.2).

Table 5.1. Samples used in this study and characters analyzed (* Indicates sequences generated by Genner *et al.* 2010).

Location	Species	Number of samples (morphometrics)	Number of sequences	Phenotype	Genbank accessions
Boadzulu Island	*M. zebra*	21	14	barred	KC208919 - KC208932
Mazinzi Reef	*M. zebra*	26	26	barred	KC208879 - KC208904
Mazinzi Reef	*M. benetos*	22	29	non-barred	KC208850 - KC208878
Illala Gap	*M. zebra*	27	19	barred	KC208974 - KC208992
Otter Point	*M. zebra*	18	23	barred	KC208951 - KC208973
Chiofu Bay	*M. zebra*	25	14*	barred	GU128640 - GU128653
Chiofu Bay	*M. chrysomallos*	23	18	non-barred	KC208933 - KC208950
Nkhata Bay	*M. zebra*	21	15*	barred	GU128829 - GU128843
Nkhata Bay	*M. callainos*	10	14	non-barred	KC208905 - KC208918
Total		193	172		

Specimens were caught in gill nets while SCUBA diving and photographed using a Canon Eos 540d. After pictures were taken the standard length was measured, specimens were fin clipped, and released back at the original collection site. As a result of past studies, which showed clear effects of sex and allometry on body shape (Herler *et al*., 2010; Kerschbaumer & Sturmbauer, 2011), we focused our analyses adult males.

Geometric Morphometrics and Statistical Analyses

We quantified body shape variation in different *Maylandia* populations using geometric morphometric analyses (Adams *et al*., 2004). Lateral pictures of individual fish

were imported into tpsDig v.2.16 (Rohlf, 2006). On each picture 16 homologous landmarks were identified and marked (Fig. 5.3, see figure caption for a description of the landmarks).

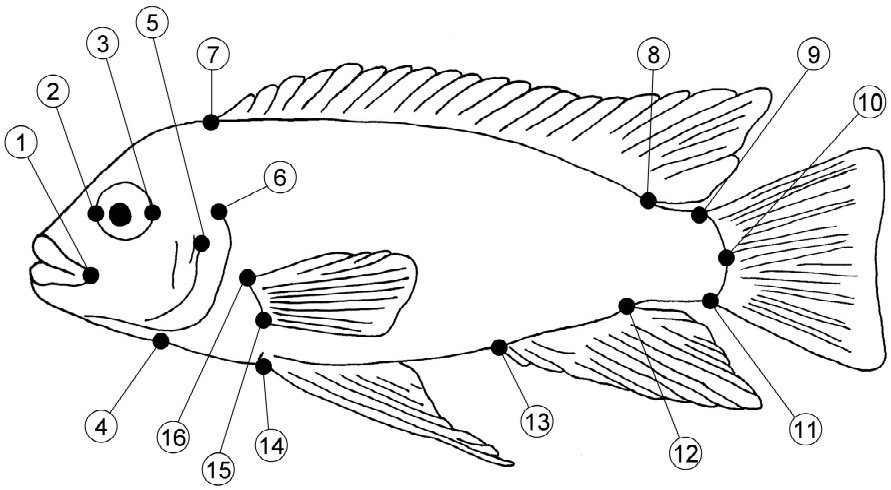

Figure 5.3. The 16 landmarks analyzed in this study: 1) most posterior point of the lips, 2) anterior edge of the eye, 3) posterior edge of the eye, 4) ventral tip of cleithrum, 5) dorsal end of pre-opercular groove, 6) dorsal origin of operculum, 7) anterior insertion of dorsal fin, 8) posterior insertion of dorsal fin, 9) upper insertion of caudal fin, 10) midpoint of the origin of caudal fin, 11) lower insertion of caudal fin, 12) posterior insertion of anal fin, 13) anterior insertion of anal fin, 14) anterior insertion of pelvic fin, 15) ventral insertion of pectoral fin, and 16) dorsal insertion of pelvic fin.

To quantify the ecological differentiation within and between *Maylandia* species, we conducted our analyses on two separate datasets. The first dataset includes species pairs from locations at which *M. zebra* and non-barred species were sympatric; the second dataset included all *M. zebra* populations. For each dataset, landmark coordinates were aligned using least-square superimposition as implemented in the program tpsRelw (Rohlf, 2007) to remove effects of translation, rotation, and scale. Based on the aligned coordinates, we generated a weight matrix consisting of partial warp scores with uniform components for each individual. To reduce data dimensionality, we subjected the weight

matrices to a principal component analysis based on the covariance matrix of the morphometric data. Unless otherwise stated, all statistical analyses were performed using SPSS 20 (IBM Inc.).

With the data collected from *M. zebra* individuals and their sympatric non-barred congeners, we wanted to address two questions: (1) is there shared site-specific variation in the barred and non-barred forms that could represent adaptation to the local environmental and (2) do barred and non-barred forms consistently differ from each other, indicating predictable differentiation at each locality? To do so, individual PC axes scores were used as dependent variables in a multivariate analysis of covariance (MANCOVA). Assumptions of multivariate normal error and homogeneity of variances and covariances were met for all analyses performed. *F*-values were approximated using Wilks' lambda and effect strengths by use of partial eta squared (η_p^2). We included species (barred vs. non-barred) and sampling site as independent variables, and used the standard length as a covariate to control for multivariate allometry. To visualize variation between species and across sites, we calculated divergence scores for each individual based on species (barred vs. non-barred) and site divergence vectors as defined by Langerhans (2009). Individual scores were then used as independent variables in tpsRegression (Rohlf, 2005) to generate thin-plate spline deformation grids highlighting shape differences among groups. Variation in divergence scores was further studied using a mixed-model nested analysis of covariance (ANCOVA) to scrutinize the results of the MANCOVA (Langerhans, 2009).

With the complete *M. zebra* data, we wanted to test the hypothesis whether 'sympatric' and 'allopatric' *M. zebra* populations consistently differed in body shape.

Again, individual PC axes scores were used as dependent variables in a multivariate analysis of covariance (MANCOVA). We included 'sympatric' vs. 'allopatric' and sampling site (nested within 'sympatric' vs. 'allopatric') as independent variables, and used the standard length as a covariate to control for multivariate allometry. As with the previous data set we calculated divergence scores for each individual based on the 'sympatric' vs. 'allopatric' and site divergence vectors as defined by Langerhans (2009). These scores were used for visualization. However, since random nested factors are not applicable for MANCOVAs and the use of fixed effects can inflate type I error rates when nested terms are significant, we also investigated variation in divergence scores using ANCOVA to scrutinize the results of the MANCOVA (Langerhans, 2009).

To quantify the effects of geographic distance and gene flow on phenotypic divergence, we conducted partial Mantel tests. To wit, we calculated pairwise phenotypic distances by first removing the effects of size using a preparatory MANCOVA and used the resulting residuals to calculate pairwise Euclidean distances between all multivariate population means (for details see Tobler & Carson, 2010). Pairwise phenotypic distances were then used as dependent variables for partial Mantel tests with 10,000 randomization, as implemented in FSTAT (Goudet, 1995). Predictor matrices for the comparison between barred and non-barred sympatric populations included colour type (same or different), geographic distance (log-transformed in kilometres), and pairwise genetic distances (Φ_{st} -values from the analyses described below). Predictor matrices for the comparison between 'allopatric' and 'sympatric' barred populations included coexistence type ('sympatric' or 'allopatric'), geographic distance (log-transformed in kilometres), and pairwise genetic distances (Φ_{st}-values from the analyses described below).

Molecular Analyses

Fish were fin clipped in the field and the tissue was either dried or stored in a preservative (20 % DMSO, 0.25 M EDTA, saturated with NaCl, pH = 7.5) until further processing. DNA was isolated using the Qiagen DNeasy blood and tissue kit (Qiagen, Hilden, Germany) following the manufacturers protocol for tissue samples. We amplified a 442 bp fragment of the mitochondrial DLoop for 143 specimens using the forward primer HapThr-2 þ 4: 5' CCTACTCCCAAAGCTAGGATC'3 and the reverse primer Fish12s: 5'TGCGGAGACTTGCATGTGTAAG'3 (Joyce *et al*., 2005). PCR was performed using the following setup: 12.2 µl of diH_2O, 2 µl of 10x PCR buffer (reaction concentration 1x), 1.6 µl of dNTP mixture (0.2 µM each, Finnzymes, Vantaa, Finland), 0.2 µl of DyNAzyme™ DNA Polymerase (1.2 U, Finnzymes, Vantaa, Finland), 1 µl of each primer (0.5 µM, Integrated DNA Technologies, Coralville, IA, USA) and 2 µl of DNA template (either pure extract, 1:10, 1:50 dilution) for a total volume of 20 µl. Amplification conditions were as follows: 94ºC for 3 min., followed by 30 cycles of 94ºC for 1 min., 58ºC 1 min., and 72ºC for 2 min., with a final elongation step at 72ºC for 10 min. 10 µl of the PCR product was purified using 4 µl ExoSAP-IT enzyme mix (Affymetrix, Santa Clara, CA, USA). The purified products were sequenced by the Sequencing Facility at Yale University. All sequences were deposited in Genbank (Accession # are given in Table 5.1).

Our own sequences and 29 additional sequences from two populations obtained from Genebank (Genner & Turner, 2012) were aligned using Geneious v. 6.0.3 (Drummond *et al*., 2011); base calls were visually inspected. General statistics of sequence variation were calculated with DnaSP v.5.10 (Librado & Rozas, 2009).

Estimates of genetic differentiation between populations were calculated as Φ_{st} with Arlequin v. 3.5.1.2 (Excoffier & Lischer, 2010) and were tested for significance using 100 permutations. A haplotype network was constructed using TCS v. 1.21 (Clement *et al*., 2000) with a connection limit of 95 %. Gaps were treated as fifth state. Relationships between populations were reconstructed using a Bayesian approach as implemented in *BEAST (Heled & Drummond, 2010) which is part of the BEAST package v.1.6.1 (Drummond & Rambaut, 2007). For the Bayesian analysis, populations were predefined according to location and phenotype. We used the GTR+I+G substitution model as determined by JModeltest (Posada, 2008). The tree prior was defined as a Yule process. We ran the MCMC simulation for 100 million generations and discarded a burn-in of 10 %. The results were checked for convergence in Tracer v. 1.5 (Rambaut & Drummond, 2009) and visualized with Densitree v. 2.01 (Bouckaert, 2010). A consensus tree was generated using TreeAnnotator v.1.6.1 discarding a burn-in of 10 % (implemented in the BEAST package v.1.6.1).

We tested for the convergence or common ancestry of the coloration phenotypes using a Bayes factors comparison (Kass & Raftery, 1995; Marek & Bond, 2006). Bayes factors analyses differs from more traditional hypotheses testing methods in not offering criteria for the absolute rejection of a null hypotheses, but instead this method allows for the evaluation of alternative hypotheses in comparison to the null hypothesis (Kass & Raftery, 1995). We constrained our phylogeny to the topologies reflecting the hypothesis of common ancestry and ran *BEAST with these constrained clades. We defined the topology generated by an unconstrained run as the null hypothesis, with populations being predefined as units. The log-files from these analyses were then used as input for

Tracer v.1.5.0, which compares the likelihood scores of the alternative hypothesis with the unconstrained tree, and generates Bayes factors. Bayes factors above 10 are considered unsupportive of hypothesized topologies, whereas Bayes factors lower than 10 are considered supportive for the respective topology (Marek & Bond, 2006). The Bayes factor approach has been employed in a variety of phylogenetic studies and represents a valuable tool when different hypotheses need to be compared (Genner & Turner, 2012; Husemann *et al*., 2012).

Results

Morphological Analyses

The PCA contrasting barred and non-barred sympatric congeners yielded eight axes describing a cumulative variance of 75 % in body shape. MANCOVA analysis showed significant effects of sampling site ($F = 8.094$, d.f. = 16, $p < 0.001$, $\eta^2_p = 0.364$) and species ($F = 28.595$, d.f. = 8, $p < 0.001$, $\eta^2_p = 0.669$) as well as an interaction of both ($F = 8.379$, d.f. = 16, $p < 0.001$, $\eta^2_p = 0.372$, Table 5.2). The ANCOVA on the species divergence scores yielded significant effects for species ($F = 48.810$, d.f. = 1, $p = 0.019$, $\eta^2_p = 0.960$) and a species by site interaction ($F = 4.561$, d.f. = 2, $p = 0.012$, $\eta^2_p = 0.071$, Table 2).

A clear pattern emerges when plotting the divergence vectors for species versus site (Fig. 5.4): Along the species divergence vector axis, congeners can clearly be distinguished, independent of their sampling site: the non-barred species are more fusiform with a lower body depth relative to *M. zebra*.

Table 5.2. Results of the multivariate analysis of covariance (MANCOVA) and the analysis of covariance (ANCOVA) on the species scores performed to test for body shape differences between color phenotypes in sympatric species pairs of *Maylandia* (species score) and between sites (sites score). *F*-ratios were approximated using Wilks' lambda, effect sizes were estimated using Partial Eta squared (η^2_p).

Effect	*F*	Hypothesis d.f.	Error d.f.	p-value	η^2_p
MANCOVA					
Intercept	5.075	8	113	<0.001	0.264
Standard length	4.893	8	113	<0.001	0.257
Species	28.595	8	113	<0.001	0.669
Site	8.094	16	226	<0.001	0.364
Species*Site	8.379	16	226	<0.001	0.372
ANCOVA species score					
Intercept	2.976	1	117.9	0.087	0.025
Standard length	2.335	1	120.0	0.129	0.019
Species	48.81	1	2.0	0.019	0.960
Site	1.317	2	2.1	0.427	0.559
Species*Site	4.561	2	120.0	0.012	0.071

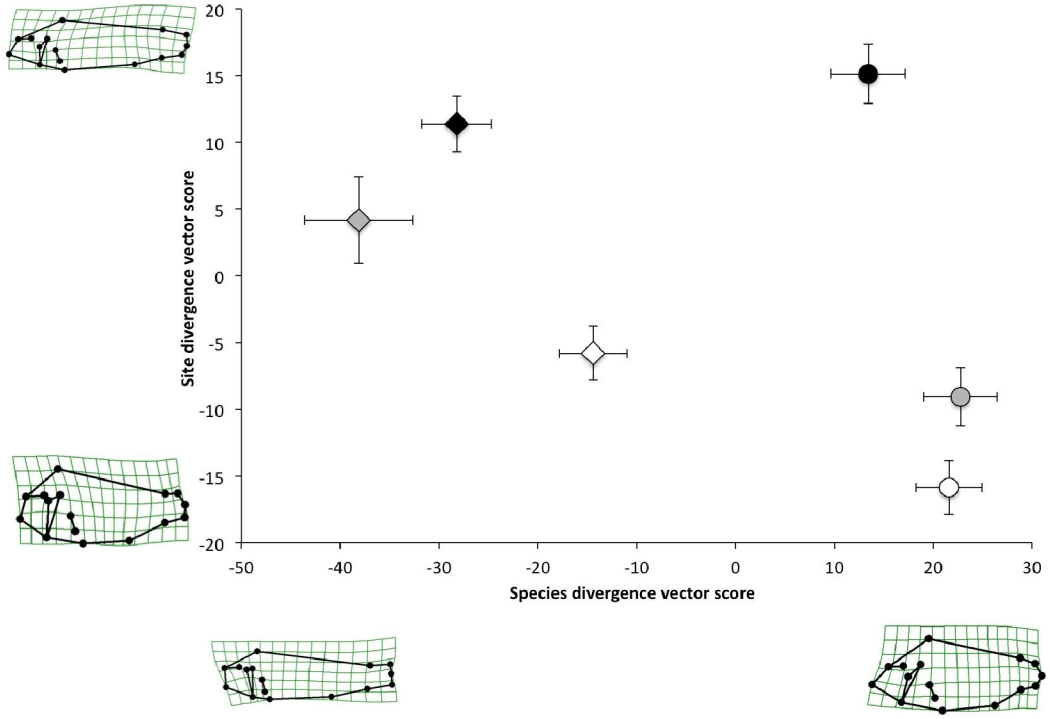

Figure 5.4. Shape variation among species and sites. Color-coding indicates different sites: Chiofu (white), Nkhata (gray), and Mazinzi (black). Symbols represent different phenotypes: barred (circles) and non-barred (diamonds).

Strong differences can also be seen in the shape of the head, the orientation of the mouth and the height of the caudal peduncle. In addition, populations with similar coloration can be distinguished along the sampling site axis. On this axis, body depth is the primary difference between members of different communities. For example, both *M. zebra* and *M. benetos* from Mazinzi Reef have the lowest body depth whereas *M. zebra* and *M. chrysomallos* individuals from Chiofu Bay are the stoutest.

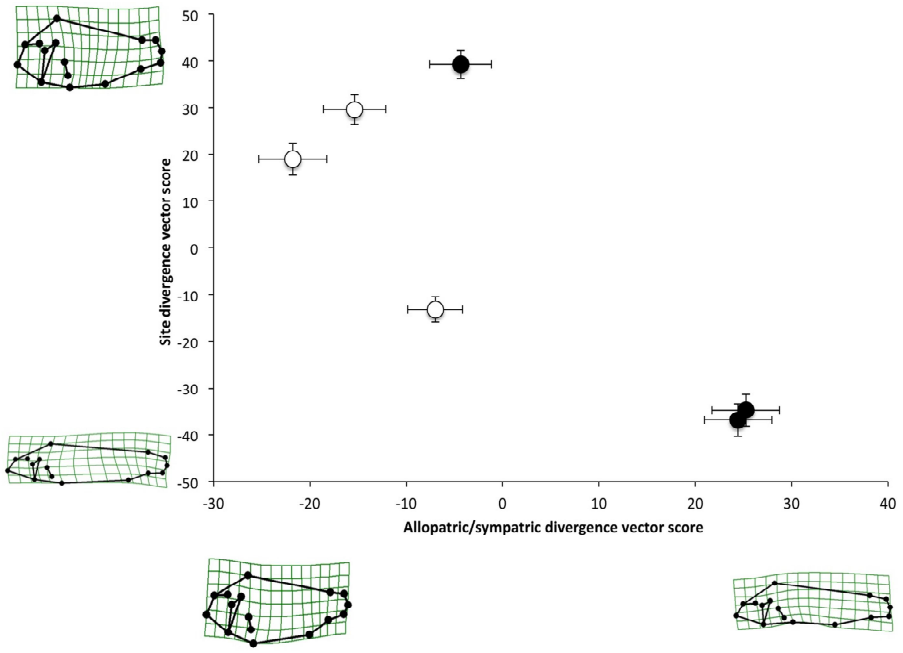

Figure 5.5. Shape variation along the 'sympatric'/'allopatric' axis and among sites; black symbols indicate 'allopatric' populations of *Maylandia zebra*, whereas white symbols stand for populations of *M. zebra* being in sympatry with a non-barred congener.

Next, we examined the effect of a non-barred congener on *M. zebra* populations. The PCA on this data set resulted in nine axes explaining a cumulative variance of 80 % in body shape. MANCOVA analysis yielded significant results (Table 3), for both the coexistence status ('sympatric' / 'allopatric') of a population ($F = 13.300$, d.f. = 9, $p <$

0.001, $\eta^2_p = 0.493$), as well as for the site effect (F = 10.332, d.f. = 36, p < 0.001, $\eta^2_p = 0.425$). The ANCOVA on the coexistence status divergence vector indicated that the presence/absence of a non-barred congener (F = 10.995, d.f. = 1, p = 0.027, $\eta^2_p = 0.723$) and site (F = 11.631, d.f. = 4, p < 0.001, $\eta^2_p = 0.262$) had a significant influence on body shape, whereas the ANCOVA on the site divergence vector scores indicated that site effects were significant (F = 83.740, d.f. = 4, p < 0.001, $\eta^2_p = 0.719$) while the coexistence status ('sympatric' / 'allopatric') was not (F = 0.943, d.f. = 1, p = 0.386).

The plot of morphological divergence along the site vector scores and the 'allopatric' / 'sympatric' vector scores (Fig. 5.5) yielded some degree of differentiation of the 'allopatric' versus 'sympatric' populations. Populations of *M. zebra* that are sympatric with a non-barred congener generally have a deeper body, whereas populations without the congener can have rather low body depth. The head shape, however, appears rather similar independent of the presence or absence of the congener.

The results from the Mantel tests analyzing the sympatric species pairs indicated that similarity in body shape is not correlated to genetic distance (Φ_{st}), geographic distance, and color phenotype. Moreover, color phenotype and geographic distance do not explain any variation in Φ_{st}. For the comparison of the six *M. zebra* populations, genetic distance significantly predicts morphological similarity (closely related populations are more similar, $r^2 = 0.74$). In addition, *M. zebra* populations in sympatry with a non-barred species are more closely related to each other than they are to populations without a non-barred species. Likewise, populations without a non-barred member appear more closely related to each other.

Table 5.3. Results of the multivariate analysis of covariance (MANCOVA) and the analysis of covariance (ANCOVA) on the species scores performed to examine body shape differences between 'allopatric' and 'sympatric' populations of *Maylandia zebra* and between sites. *F*-ratios were approximated using Wilks' lambda, effect sizes were estimated using Partial Eta squared (η^2_p).

Effect	*F*	Hypothesis d.f.	Error d.f.	p-value.	η^2_p
MANCOVA					
Intercept	4.501	9.0	123.0	<0.001	0.248
Standard length	4.518	9.0	123.0	<0.001	0.248
'Sympatric' / 'Allopatric'	13.300	9.0	123.0	<0.001	0.493
Site ('Sympatric' / 'Allopatric')	10.332	36.0	462.7	<0.001	0.425
ANCOVA 'sympatric' / 'allopatric' score					
Intercept	29.140	1.0	126.0	<0.001	0.188
Standard length	31.123	1.0	131.0	<0.001	0.192
'Sympatric' / 'Allopatric'	10.995	1.0	4.2	0.027	0.723
Site ('Sympatric' / 'Allopatric')	11.631	4.0	131.0	<0.001	0.262
ANCOVA site score					
Intercept	14.963	1.0	27.1	0.001	0.356
Standard length	23.835	1.0	131.0	<0.001	0.154
'Sympatric' / 'Allopatric'	0.943	1.0	4.0	0.386	0.190
Site ('Sympatric' / 'Allopatric')	83.740	4.0	131.0	<0.001	0.719

Molecular Analyses

We sequenced a 442 bp fragment of the mitochondrial DLoop for 143 specimens and obtained 29 additional sequences from a study by Genner & Turner (2012) for a total of 172 sequences belonging to 9 populations of *Maylandia* (Table 1). The alignment contained 5 gaps and 25 variable sites, 18 of which were parsimony informative. A total of 28 haplotypes were recovered. The total nucleotide diversity was 0.00604. Φ_{st} estimates were generally high and ranged from 0.02573 (*M. zebra* from Boadzulu Island and *M. chrysomallos* from Chiofu Bay) to 0.93354 (between *M. benetos* from Mazinzi Reef and *M. zebra* from Otter Point). All comparisons except for the one, between *M. zebra* from Boadzulu Island and *M. chrysomallos* from Chiofu Bay, were significant

(Table 4). Geographic distance did not explain the extent of genetic differentiation between populations (Z = 815.8484, r = -0.1786, one-sided p = 0.7752).

The haplotype network showed no clear structure (Fig. 5.6). All haplotypes were closely related and no more than three mutational steps separated any haplotype. *Maylandia benetos* from Mazinzi Reef was the least diverse population with only a single haplotype, while *M. zebra* was the most diverse with eight haplotypes. Across all species, three haplotypes were especially common. The first of these common haplotypes was shared only between individuals from Illala Gap and Otter Point, while the second common haplotype was found in *M. benetos* and *M. zebra* from Mazinzi Reef and *M. callainos* from Nkhata Bay. The third common haplotype was most widely distributed and was found in *M. callainos* and *M. zebra* from Nkhata Bay, *M. chrysomallos* from Chiofu Bay, and *M. zebra* from Boadzulu Island, Otter Point, and Illala Gap. A fourth less common haplotype was shared between *M. callainos* from Nkhata Bay and *M. zebra* from Illala Gap. All other haplotypes were location specific.

Our efforts to reconstruct the relationships among populations and species yielded a tree with low support (Appendix 5.1). However, the tree shows one well supported major split (pp = 100) into two main branches, which group *M. benetos* and *M zebra* from Mazinzi Reef together with *M. callainos* from Nkhata Bay and separates these three populations from all others. The only other well supported node (pp = 94) groups the *M. zebra* populations from Illala Gap and Otter Point together, which represent the two geographically closest locations. All other branches have very low support and the relationships cannot be regarded as reliable. In addition, we reanalysed AFLP data provided by Allender *et al*. (2003), which included *M. benetos* and *M. zebra* from

Mazinzi Reef, and *M. callainos* from Nkhata Bay, as well as several other *Maylandia* species.

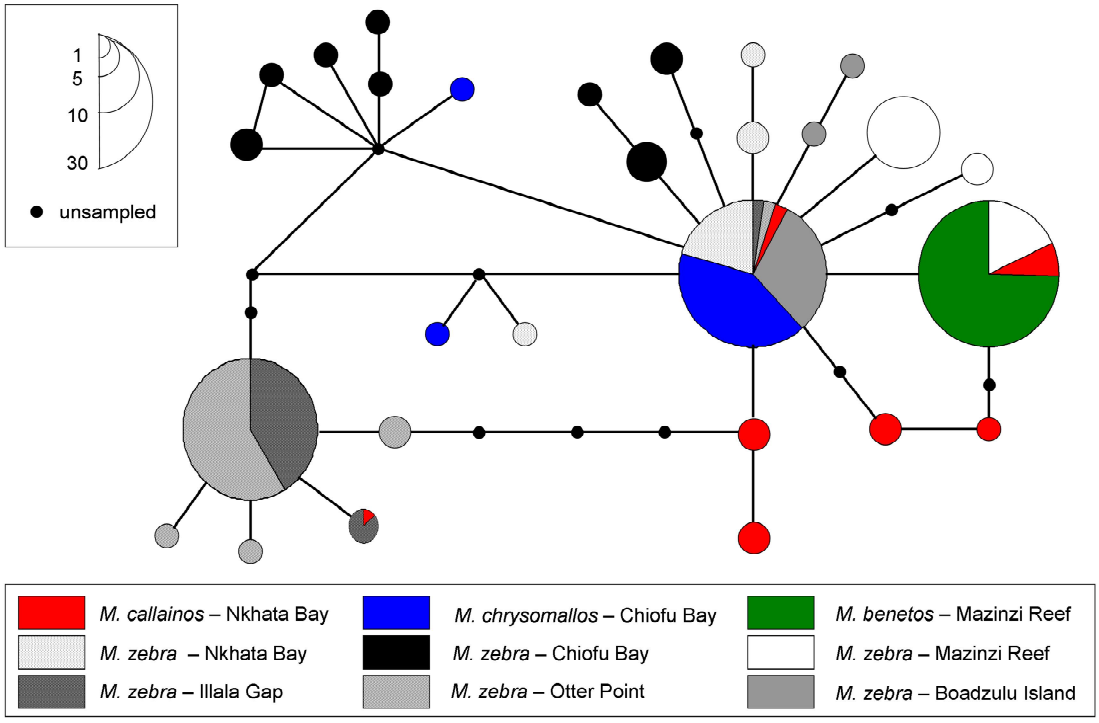

Figure 5.6. Statistical parsimony network for the investigated *Maylandia* populations with haplotypes connected at a 95 % significance level.

Table 5.4. Φ_{ST} estimates of genetic differentiation generated from D-Loop sequences for all populations and species with Arlequin. Significance ($\alpha = 0.05$) as determined by 100 permutations is indicated by an asterisk; Mb – *M. benetos*, Mz – *M. zebra*, Mcal – *M. callainos*, Mchr – *M. chrysomallos*.

Pop.	Mb – Mazinzi	Mz – Mazinzi	Mcal – Nkhata	Mz – Boadzulu	Mchr – Chiofu	Mz – Chiofu	Mz – Nkhata	Mz – Otter
Mb – Mazinzi	0							
Mz – Mazinzi	0.587*	0						
Mcal – Nkhata	0.556*	0.284*	0					
Mz – Boadzulu	0.924*	0.288*	0.244*	0				
Mchr – Chiofu	0.856*	0.284*	0.249*	0.026	0			
Mz – Chiofu	0.655*	0.323*	0.254*	0.229*	0.214*	0		
Mz – Nkhata	0.797*	0.300*	0.248*	0.167*	0.127*	0.237*	0	
Mz – Otter	0.934*	0.801*	0.761*	0.866*	0.848*	0.708*	0.817*	0
Mz – Illala	0.932*	0.790*	0.747*	0.857*	0.840*	0.686*	0.804*	0.091*

We applied Bayesian and Neighbornet phylogenetic methods to the AFLP data. Neither approach provided good resolution, yet both indicated that the species with the non-barred phenotype are not monophyletic and species with different coloration phenotypes group together (Appendix 5.2, also see Allender *et al*., 2003 for more details). Our Bayes factor analysis clearly rejected monophyly of the coloration phenotypes (Bayes factor: 941.002; H_A: clustering by coloration phenotypes -883.347 ± 0.094 (mean LnL ± SE); unconstrained tree H_0: LnL -876.5 ± 0.118).

Discussion

Theory predicts that stable coexistence is only possible if sympatric species are diverged in ecological and mating traits (Gause, 1934; Armstrong & McGehee, 1980; Dieckmann & Doebeli, 1999; Vandermeer *et al*., 2002). Though it had been hypothesized that cichlids might represent an exception to this rule (Liem, 1980), more recently evidence suggests that resource partitioning might be common in cichlids (e.g. Albertson, 2008; Arnegard, 2009; Danley, 2011). Closely related, sympatric species of Lake Malawi cichlids have been shown to partition a habitat by depth (Albertson, 2008) and microhabitat (Danley, 2011) but not diet (Martin & Genner, 2009). Thus habitat partitioning appears an important mechanism to avoid competition with congenerics in the rock-dwelling cichlids of Lake Malawi.

Building on a previous study that documented microhabitat partitioning between barred and non-barred *Maylandia* species (Danley 2011), this study examined similar replicate species pairs to determine if morphological variation can distinguish them. At each location with species pairs of different coloration phenotypes, both species were morphologically differentiated from each other. Interestingly, the direction of shape

change was similar at all locations (Fig. 5.4, Table 5.3). The non-barred species consistently had lower body depth and were more fusiform in comparison to the barred species. Such differences in body depth generally are associated with swimming performance in fish. A shallower, fusiform body allows for greater sustained swimming performance and therefore can be an advantage in open habitats, whereas a deeper body allows for higher maneuverability and is advantageous in habitats with complex structures (e.g. Barlow, 1972; Langerhans & Reznick, 2009).

This finding is consistent with previous behavioural studies. Danley (2011) found that *M. zebra* prefer cobble rich habitat patches while the non-barred *M. benetos* prefer bedrock at Mazinzi Reef. Similar preferences for simple and complex habitats are found for *M. chrysomallos* and *M. zebra,* respectively, at Chiofu Bay (P.D.D. & M.H., unpublished data). Holzberg (1978) also found differences in the preference of territory size and feeding habitat in the sympatric *M. zebra* and *M. callainos* at Nkhata Bay. In our study, it appears that congeneric sympatric species have adapted to different microhabitats with deep bodied species common in cobble rich habitats and slender bodied species common in structurally simpler habitats. The observed divergence body shape may facilitate partitioning of the habitat to minimize competition for both food and territories.

In addition to body shape, male color pattern may provide an adaptive advantage in their respective microenvironments. In Lake Malawi's rock-dwelling cichlids, coloration is thought to be cryptic in females and conspicuous in males (Roberts *et al*., 2009). However, a meta-analysis by Seehausen *et al*. (1999) found that barred patterns are generally associated with complex habitats in cichlids (e.g. rocky or vegetated) and

suggested an adaptive function of the barred pattern (but see Deutsch, 1997 for an alternate opinion). If an interrupted colour pattern is an adaptation to complex habitats, the non-barred phenotype might be more cryptic over homogeneous habitats such as bedrock. While an explicit test of the adaptive benefit of the barred and non-barred phenotypes is needed, we suggest that both coloration and body shape might be under correlational selection in response to their apparent habitat preferences (Brodie, 1992; McGlothlin *et al*., 2005; Bergstrom, 2007; Carlsbeek & Irschick, 2007; Roff & Fairbairn, 2012).

Furthermore, habitat preference is correlated with aggressive behaviour across a number of *Maylandia* species. Bedrock inhabiting species, such as the non-barred *M. benetos*, are more aggressive than the barred *M. zebra* at Mazinzi Reef (Danley 2011). Likewise, at Chiofu Bay the more aggressive *M. chrysomallos* tends to prefer small cobble and bedrock while *M. zebra* is less aggressive and prefers medium to large coble (P.D.D. & M.H. unpublished data). Our results suggest that adaptation to simple versus complex microhabitats may have lead to the divergence of a suite of characteristics including body shape characteristics (body depth, depth of the caudal peduncle, head shape), male color pattern, habitat preference and male aggressive behaviour through correlational selection.

It is unclear what evolutionary processes have driven the differences in sympatric species. One process might involved character displacement. Character displacement has driven rapid phenotypic divergence in a number of closely related, sympatric species (e.g. Adams, 2010; Scott & Johnson, 2010; Magalhaes *et al*., 2012). For example, character displacement is thought to have driven three-spined sticklebacks (Schluter & McPhail,

1992; Pritchard & Schluter, 2001) and Nicaraguan crater lake cichlids (Barluenga *et al*., 2006) into limnetic and benthic forms, beak size divergence in Galapagos finches (Schluter & Grant, 1984; Grant & Grant, 2006) and ecomorph divergence in sympatric species of *Anolis* lizards (Losos *et al*., 1998). If character displacement has driven this divergence, one would expect that populations of *M. zebra* in communities lacking a sympatric non-barred congener would occupy more morphospace than populations with a non-barred sympatric congener (Willis *et al*., 2005) which is, in fact, the pattern observed in our data (Appendix 5.3). Further, populations of *M. zebra* lacking the congener often have a lower body shape compared to populations with a congener (Fig. 5.5). While this pattern is consistent with character displacement, explicit tests using manipulative approaches are needed to conclude that character displacement caused the observed differentiation (Schluter, 2000; Stuart & Losos, 2013).

Alternatively, ecological sorting could explain the consistent phenotypic differences observed between the sympatric species pairs (Armbruster *et al*., 1994; Schluter, 2000; Sax *et al*., 2007; Dangles *et al*., 2008). Ecological sorting occurs when ecologically similar species are competitively excluded from a location. This process generates communities of ecologically divergent species capable of coexisting (Armbruster *et al*., 1994; Schluter, 2000). For example, ecological sorting has resulted in size assortment of Caribbean *Anolis* (Losos, 1990). Given the dynamic nature of Lake Malawi's water level, it is possible that ecological sorting has produced the observed phenotypic pattern. Throughout its history, Lake Malawi experienced multiple desiccation/inundation events producing cycles of admixture and separation of complex, geographically distinct cichlid communities (Danley *et al*., 2012a). The most recent

desiccation event occurred within the past 700 years and would have rendered all of our collection locations except for Nkhata Bay dry land (Owen *et al*., 1990). During any of these cycles, ecological sorting may have occurred as communities assembled. However, ecological sorting suggests that species are static units that only survive in sympatry if they are preadapted to specific niches which are not occupied by others (Armbruster *et al*., 1994). Yet, our data suggest that *M. zebra* populations have experienced site-specific shape changes as an adaptive or plastic response to local selective pressures. A future study examines the genetic and plastic components of *Maylandia* body shape variation (Husemann *et al*., in prep).

While phenotypic differentiation is most apparent between *M. zebra* and the non-barred species, this study also revealed morphological differentiation within *M. zebra*. Among East African cichlids, geographic separation and local adaptation often result in strong phenotypic and genetic differentiation (e.g. Bouton *et al*., 1999, 2002; Streelman *et al*., 2007; Postl *et al*., 2008; Kerschbaumer *et al*., 2011; Spreitzer *et al*., 2012; Magalhaes *et al*., 2012). Our results are consistent with these previous studies: *M. zebra* is morphologically differentiated for almost all of the analyzed populations. The only exceptions are the geographically close populations of Otter Point and Illala Gap whose morphological similarity is likely the result of ongoing gene flow which is supported by the extensive sharing of haplotypes (Fig. 5.6) and the low estimates of genetic divergence of these populations (Table 5.2). Among populations that are more widely separated, however, no isolation by distance is found for either the genetic or the morphological data set. This suggests that beyond a very limited scale geographic distance does not determine morphological or genetic similarity. Instead, the morphological differentiation

of *M. zebra* populations is likely the result of local selective pressures. A similar pattern of local adaptation is seen in each of the non-barred species (Fig. 5.4) and has been shown for other morphological characters in *M. zebra* (Streelman *et al*., 2007), further underlining the important effects of the local environment on ecomorphology.

While microhabitat partitioning may explain the divergence in the suite of characters associated with the barred and non-barred phenotypes, we cannot exclude the role that other factors such as sexual selection and reinforcement may play in driving this divergence. Future studies are needed to examine the role that this suite of characters plays in the reproductive behavior of cichlids (Salzburger, 2009; van Staaden & Smith, 2011).

Conclusions

In this study we show that consistent differences in an ecological trait exist in replicated species pairs of the rock-dwelling genus *Maylandia*. Morphospace occupation is smaller in sympatry than in allopatry suggesting that character displacement might play a role for divergence. Coloration and body shape appear to evolve in specific combinations likely as adaptation to specific microhabitats resulting from correlational selection. Allopatric populations of *M. zebra* are molecularly and phenotypically differentiated as result of geographic isolation and local selection. The divergence of *Maylandia* populations appears to be strongly driven by local adaptation as well as by sympatry with closely related species. The correlation of ecologically selected traits such as body shape with sexual traits may facilitate the coexistence of congers and help to generate and maintain species diversity in this system.

CHAPTER SIX

Genetic and Plastic Components of Body Shape and the Potential for Transgressive Segregation in a Pair of Closely Related Malawi Cichlids

Introduction

Understanding the drivers of phenotypic diversification remains one of the central goals of evolution. Phenotypic divergence is often driven by heritable responses to selection or by plastic responses to the environment (West-Eberhard 1989). Historically, selection has been assumed to be the prevalent force in driving phenotypic divergence among populations (Darwin 1859, Coyne & Orr 2004), while the role of phenotypic plasticity in biological diversification remains controversial (Pfennig et al. 2010). Increasingly, however, phenotypic plasticity is being recognized as a potential mechanism promoting divergence by generating new phenotypes selection may act upon (Pfennig et al. 2010). Transgressive segregation is an additional mechanism for the generation of phenotypic variation (Rieseberg et al. 1999, Seehausen 2004). Transgressive segregation occurs when hybrid phenotypes exceed the phenotypic distribution of the parental species (Rieseberg et al. 1999). Selection on the newly generated phenotypic variation can allow these individuals with unique trait combinations to occupy novel niches (Via & Lande 1985, Seehausen 2004, Albertson & Kocher 2005, Ghalambor et al. 2007). The potential for transgressive segregation is determined by the genetic architecture of a phenotype (Rieseberg et al. 1999, Albertson & Kocher 2005). Complementary gene action with antagonistic gene effects is thought to be the most important genetic mechanism underlying transgressive segregation. Therefore, traits

lacking antagonistic alleles due to consistent directional selection are thought to exhibit little or no transgressive segregation (Rieseberg et al. 1999, Albertson & Kocher 2005). Additionally, the degree of species divergence is thought to be positively correlated with transgressive segregation, with hybrids between more distantly related species exhibiting more transgressive phenotypes (Rieseberg et al. 1999, Stelkens & Seehausen 2009, Stelkens et al. 2009). Due to their potential to generate phenotypic variation both phenotypic plasticity and transgressive segregation have been suggested to have played a role in rapid species radiations (Stauffer & van Snik Gray 2004, Seehausen 2004, Bell & Travis 2005, Albertson & Kocher 2005, Herder et al. 2006, Kerschbaumer et al. 2011, Genner & Turner 2012).

With more than 2000 species in the three East African Great Lakes (Tanganyika, Victoria, and Malawi), East African cichlids represent the most diverse extant vertebrate radiation known. East African cichlids exhibit an extraordinary amount of phenotypic diversity allowing cichlids to occupy all major ecological niches within the lakes (Seehausen 2006, Sturmbauer et al. 2011). Of the three lakes, Lake Malawi harbors the most species rich radiation with more than 700 species (Kocher 2004, Danley et al. 2012a). Selection is thought to be the main driver of diversification in all stages of the radiation (Streelman & Danley 2003). At the first and second stage natural selection led to macrohabitat divergence and the differentiation trophic groups into genera, respectively. During the most recent stage of their diversification natural and sexual selection drove the divergence of signaling phenotypes, microhabitat preferences, and body morphologies (Sturmbauer 1998, Danley & Kocher 2001, Streelman & Danley 2003, Streelman et al. 2007, Danley et al. 2012b, Husemann et al. in review).

Here, we examine the divergence of an ecologically important character, body shape, between closely related species. Body shape has been shown to evolve in response to ecological selection in fish and therefore can be used as an ecological marker when studying differentiation in natural populations (Tobler et al. 2008, Langerhans 2009, Odhiambo et al. 2011). Body shape differentiation of fish has been documented in response to a variety of different sources of selection including predation (Langerhans 2009), abiotic environmental factors (Neves & Monteiro 2003, Tobler et al. 2008, Kerschbaumer et al. 2011), and sympatry with close relatives and the resulting competition (Scott & Johnson 2010, Husemann et al. in review).

The divergence of body shape between and within closely related species is well documented in the rock-dwelling genus *Maylandia*. With 31 species, *Maylandia* is among the most species rich genera in Lake Malawi (Stauffer et al. 2013). Species within this genus are distinguished by differences in male mating coloration, slight differences in body shape, and behavioral reproductive isolation when sympatric (Stauffer et al. 1997, 2013, Kidd et al 2006, Plenderleith et al. 2005). Within species, population level differences in body shape have been observed (Streelman et al. 2007), and such differences have been attributed to local selection pressures as well as heterospecific resource competition (Husemann et al. in review). However, whether these shape differences are heritable or the result of phenotypic plasticity remains unknown.

In order to address this question we performed a common garden experiment and generated hybrid crosses between two closely related East African cichlid species. Previous studies have indicated that body shape in cichlids and other fish has a strong plastic component, yet most studies also found that some shape components are heritable

(Kerschbaumer et al. 2011, Magalhaes et al. 2012, McCairns & Bernatchez 2012). We expected to observe a similar pattern in which at least part of the differences between species would be conserved in the individuals raised under standardized conditions. Our experimental design further allowed us to investigate the mode of gene action underlying the differences in body shape and to determine the potential for transgressive segregation. As body shape is a complex, modular phenotype, we expected that the differences would not be explained by a simple additive model but would rather involve epistatic interactions, as it has been shown for sticklebacks (Schluter et al. 2004). Complex gene interactions are thought to promote transgressive segregation and body shape is prone to exhibit transgressive segregation (Burke & Arnold 2001, Rieseberg et al. 1999, 2003a, Bell & Travis 2005). However, the amount of transgression exhibited in a cross is often correlated with the genetic distance between the parentals (Stelkens & Seehausen 2009, Stelkens et al. 2009). As the two study species are very close relatives, we expected to observe rather limited amounts of transgression in our cross (Albertson et al. 1999, Allender et al. 2003).

Material and Methods

Study Species

We used two closely related species of the rock-dwelling cichlid genus *Maylandia* to study the plastic and genetic components of body shape differentiation and test for transgressive segregation. *Maylandia benetos* is a microendemic; it occurs at a single location in the lake, Mazinzi Reef, where it is sympatric with two other *Maylandia* species, including *Maylandia zebra*. *Maylandia zebra* is one of the few rock-dwelling

cichlids (mbuna) found at most rocky locations throughout the lake (Ribbink et al. 1983, Ding et al. in review). A previous study has shown that sympatric barred and non-barred *Maylandia* species, including *M. benetos* (non-barred) and *M. zebra* (barred) from Mazinzi Reef, are differentiated in their body shape in a predictable manner based on their coloration phenotype (Husemann et al. in review). To further understand the repeated, parallel divergence of body shape differentiating barred and non-barred species, we used *M. benetos* and *M. zebra* as a model to study the transmission of body shape variation. Specifically, 1) populations of both were raised in identical aquaculture environments to quantify the degree of plasticity influencing this phenotype, and 2) these species were crossed in the lab to investigate the underlying mode of gene action and the amount of transgressive segregation observed in the body shape phenotype.

Sampling

We collected adult specimens of *M. zebra* (N=38) and *M. benetos* (N=44) in the summers of 2010 and 2012 at Mazinzi Reef. Specimens were caught in nets while using SCUBA and photographed using a Canon Eos 540d.

Table 6.1. Sampling list. The table shows the number of sampled males and females, the total number of individuals used in the study, and the rearing environment for each group.

Species	Environment	# males	# females	total
M. benetos	Field	32	12	44
M. zebra	Field	33	5	38
M. benetos	Lab	38	17	55
M. zebra	Lab	57	24	81
F1	Lab	49	47	96
F2	Lab	117	209	326
Backcross to *M. zebra*	Lab	22	20	42
Backcross to *M. benetos*	Lab	15	5	20
total		363	339	702

In addition, we analyzed *M. zebra* (N=81) and *M. benetos* (N=55) descended from wild caught populations that have been maintained as laboratory stocks for approximately 12 generations. Fish were maintained in 110 cm × 28 cm × 30 cm tanks and fed with high quality flake food. Bidirectional F1 (N=96) were produced, and these F1 were used to produce F2 (N=326). We then generated backcrosses to both parental species (F1 x *M. benetos* N= 20, F1 x *M. zebra* N= 42). A total of 702 individuals were used in this study (Table 1). Pictures were taken with a Canon Eos 540d under standardized conditions including a ruler as length standard. All specimens were sexed and the standard length was measured.

Morphometric Analyses

We quantified body shape variation in the two species and hybrid generations using geometric morphometric analyses (Adams et al. 2004). Lateral pictures of individual fish were imported into tpsDig v.2.16 (Rohlf 2006), and we digitized 16 landmarks (Fig. 6.1, see figure caption for a description of the landmarks). To address our question regarding the genetic and environmental components of body shape we used data collected from wild caught and lab reared *M. benetos* and *M. zebra*. In a second analysis we tried to determine the mode of gene action and test for transgressive segregation in the hybrid generations. For this we included all laboratory reared parental generations and the F1, F2, and backcrosses.

For each dataset, landmark coordinates were aligned using least-square superimposition as implemented in the program tpsRelw (Rohlf, 2007) to remove effects of translation, rotation, and scaling. Unless otherwise stated, all statistical analyses were performed using SPSS 20 (IBM Inc.).

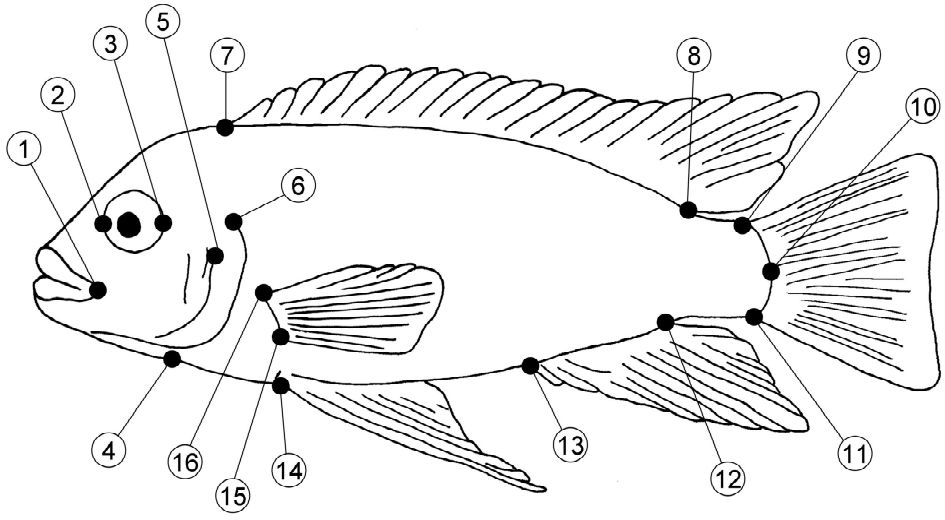

Figure 6.1. The 16 landmarks analyzed in this study:1) most posterior point of the lips, 2) anterior edge of the eye, 3) posterior edge of the eye, 4) ventral tip of cleithrum, 5) dorsal end of pre-opercular groove, 6) dorsal origin of operculum, 7) anterior insertion of dorsal fin, 8) posterior insertion of dorsal fin, 9) upper insertion of caudal fin, 10) midpoint of the origin of caudal fin, 11) lower insertion of caudal fin, 12) posterior insertion of anal fin, 13) anterior insertion of anal fin, 14) anterior insertion of pelvic fin, 15) ventral insertion of pectoral fin, and 16) dorsal insertion of pelvic fin.

Estimating Genetic and Plastic Components

To distinguish genetic and plastic components of shape differentiation we analyzed the similarities and differences of body shape in wild caught and lab reared populations of *M. zebra* and *M. benetos* from Mazinzi Reef. A total of 218 Individuals were included in this analysis. Based on the aligned landmark coordinates and the consensus configuration, we generated a weight matrix by calculating partial warp scores with uniform components for each individual. To reduce the data complexity we subjected the weight matrix to a principal component analysis based on the covariance matrix of all landmarks to generate a relative warp matrix. Standard length and sex were included in the analyses as covariates to control for multivariate allometry. To visualize

variation between laboratory and field environment and species, we calculated divergence scores for each individual based on species and environment divergence vectors as defined by Langerhans (2009). Individual scores were then used as independent variables in tpsRegression (Rohlf 2005) to generate thin-plate spline deformation grids highlighting shape differences among groups (Zelditch et al. 2004).

Further we performed a multivariate analysis of covariance (MANCOVA) to compare laboratory and field samples of the parental species. We used standard length as a covariate, and gender, environment (laboratory vs. field) and species (*M. zebra* vs. *M. benetos*) as independent variables. This approach can be used estimate the effects of phenotypic plasticity and genetic components on shape: if all observed differences between the species were the results of environmental plasticity one would expect that these differences disappear in a common rearing environment. If, however, differences have a genetic component species differences should be maintained and no differences should be seen between rearing environments (Tobler et al. 2008, Kerschbaumer et al. 2011, McCairns & Bernatchez 2012, Torres-Dowdall et al. 2012).

Analyses of Body Shape in Laboratory Crosses

To investigate the mode of gene action and test for transgressive segregation in body shape, we examined the laboratory reared parental species and their F1, F2, and backcross hybrids for a total of 620 individuals. The data were processed as described above and the means and variances for each generation's divergence score were calculated and plotted. If the body shape variation between the parental species is heritable and the genetic basis for these differences is additive, all hybrid generations' mean phenotypes are expected to be intermediate to the parentals. However, the variance

in the F2 generation is expected to be higher in comparison to the parentals and the F1 hybrid generation (Barson et al. 2007). To test for additive effects, we compared the additive model of gene action to the additive-dominance model using the joint-scaling test to determine which model better fits our data (Lynch & Walsh 1998). The joint-scaling method can also be used to detect the action of epistasis, however testing for epistasis would require more than the six lines available from our cross (P1, P2, F1, F2, BC1, BC2). Therefore, we used a t-test based on P1, P2, F1 and F2 data as suggested by Lynch and Walsh (1998) to evaluate epistasis.

To estimate and quantify the amount of transgressive segregation found in our cross we used two separate approaches. First we employed the method developed by Stelkens et al. (2009): we determined the range for the combined parentals and for the complete data set for each PC axis. The amount of transgression occurring along each axis was then calculated by subtracting the range of the parentals from the total range of the data set along the axis. The difference between the parental range and the total range was then divided by the range of the parentals. The total amount of transgression occurring in the hybrid generations (F1, F2 and backcrosses) was then calculated by summing up the transgression found on each axis adjusted for the percent of variance explained by the axis (Stelkens et al. 2009).

To parallel Stelkens et al.'s (2009) study of transgressive segregation and genetic distance in cichlids, we calculated the genetic p-distance for the two species using 163 mitochondrial D-Loop sequences from both species (*M. benetos* N = 85, *M. zebra* N = 78) provided from previous studies (Husemann et al. in review a, b). Genetic distances were calculated in MEGA5 (Tamura et al. 2011).

As sample sizes differed across generations, we performed a second analysis testing for transgression while adjusting for different sample sizes. We generated estimates of convex hull volumes for each parental species, the combined parentals, and the F2 generation. We removed variation due to sex and size by using the residuals from a preparatory MANCOVA. We performed a PC analysis and used the first 9 axes to calculate a convex hull for each group using the Quickhull algorithm (Barber et al. 1999). The convex hull of a set of points is a geometric measure describing the smallest convex set that contains all points in that dataset (Barber et al. 1999). Due to different sample sizes among groups we used a randomization procedure to calculate morphospace as described in Tobler & Carson (2010). Random distributions of morphospace were generated using 1000 iterations of randomly selected specimens with replacement from the respective pool of individuals. A convex hull was calculated for each sample. Means and confidence intervals were calculated for each group through the examination of 1,000 iterations of this process. If body size exhibits transgressive segregation the F2 is expected to occupy significantly more morphospace than that of the combined parental generations.

Results

Genetic and Plastic Components of Body Shape

The analysis of wild caught and laboratory reared fish revealed significant plastic and genetic components that influence body shape. The greatest differences between the species, independent of the environment, were found in the shape and slope of the head and body depth (Fig. 6.2). When looking at the differences between the laboratory and

field raised fish, both species had much deeper bodies with higher caudal peduncles in the laboratory whereas body shape was more elongated and fusiform in the wild caught individuals. The variance in the species score was higher in the laboratory compared to the wild caught samples.

Table 6.2. Results of the multivariate analysis of covariance (MANCOVA) of body shape in the laboratory – field comparison of *M. benetos* and *M. zebra*. F-ratios were approximated using Wilks' lambda, effect sizes were estimated with partial Eta squared (η^2_p). Significant values are printed in bold.

Effect	Wilk's Lambda	F	Hypothesis d.f.	Error d.f.	p	η^2_p	Relative variance
Standard length	0.822	4.872	9.000	202.000	**<0.001**	0.178	0.304
Environment	0.415	31.671	9.000	202.000	**<0.001**	0.585	1.000
Sex	0.858	3.713	9.000	202.000	**<0.001**	0.142	0.243
Species	0.600	14.990	9.000	202.000	**<0.001**	0.400	0.684
Environment * Species	0.644	12.409	9.000	202.000	**<0.001**	0.356	0.609
Sex * Species	0.930	1.694	9.000	202.000	0.092	0.070	0.120
Environment * Sex	0.877	3.162	9.000	202.000	**0.001**	0.123	0.210

The MANCOVA showed that all main effects and most interaction terms were significant (Table 2); however, the effects of standard length and sex and the interaction terms involving sex were generally weak ($\eta^2_p < 0.15$). The rearing environment (laboratory vs. field, plastic component) had the strongest effect on body shape ($\eta^2_p = 0.585$), followed by the species identity (genetic component, $\eta^2_p = 0.400$). The interaction of species by environment was significant as well, yet had a weaker effect ($\eta^2_p = 0.356$).

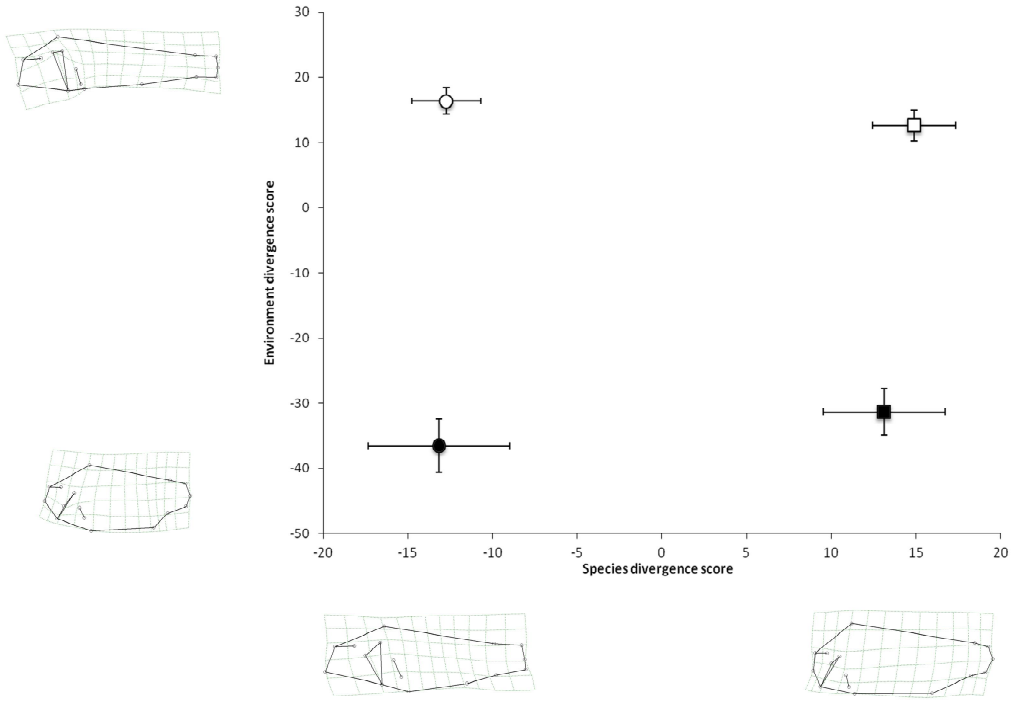

Figure 6.2. Plot of the environmental divergence and species divergence vector scores for *M. benetos* (squares) and *M. zebra* (circles) sampled in the field (black) and laboratory raised (white). Error bars represent the standard errors of the divergence vector scores.

Body Shape in Laboratory Crosses

In the laboratory-bred stocks, the means of the divergence vector scores of the parental species defined the two phenotypic extremes and the means of all hybrid generations had intermediate values relative to the parentals (Table 3). The F1 divergence score mean was strongly skewed towards *M. benetos*, whereas the F2 mean divergence score was roughly intermediate between both parentals. The backcrosses had intermediate divergence scores, though the backcross to *M. benetos* was highly skewed towards *M. benetos*.

An analysis of the mode of gene action of body shape differences via the joint-scaling test rejected additivity (p.A = 0). The rejection of the additive model can be easily

observed in a plot of the means and variances of each generation. If the phenotype follows the additive genetic model, a plot of the means versus variances for each generation would produce a triangular pattern with the parental points defining the base of the triangle and the F2 its apex (Barson et al. 2007, Fig. 6.3).

Table 6.3. Means and variances for species vector scores for cross lines.

Line	Mean	Variance
M. zebra	39.445661	637.0195
M. benetos	-32.811476	1094.2716
F1	-18.058665	943.7731
F2	2.276948	1164.6087
Backcross to *M. zebra*	5.178079	691.6697
Backcross to *M. benetos*	-30.830032	1493.0822

Our data clearly deviate from these expectations: while the parentals represented the morphological extremes (*M. benetos* N=55, μ = -32.81, σ^2 = 1094.27; *M. zebra* N=81, μ = 39.45, σ^2 = 637.02, Table 6.3), the backcross to *M. benetos* had the highest variance (N=20, μ = -30.83, σ^2 = 1493.08). Further the F1 (N=96, μ = -18.06, σ^2 = 943.7731) and both backcrosses deviated from the expected values under additivity and were skewed towards *M. benetos.* Further the variance of the F1 generation was lower than that observed in *M. benetos*. The variance of the F2 (N=326, μ = 2.28, σ^2 = 1164.61) was only slightly higher than that in *M. benetos* (Fig. 6.3).

The additive-dominant model was rejected as well (p.AD = $1.01*e^{-14}$); yet, when comparing both models the additive-dominant model better explains the data (p.A.AD = $4.68*e^{-08}$). As we did not have sufficient hybrid lines we could not use the joint-scaling test for epistasis. Instead we used the test based on the variances of the parental lines, F1 and F2 as proposed by Lynch & Walsh (1998). The t-test could not reject epistatic effects

(test statistic: -0.22; epistasis can only be rejected if this value is above 1.96; Lynch & Walsh 1998, Albertson et al. 2003a).

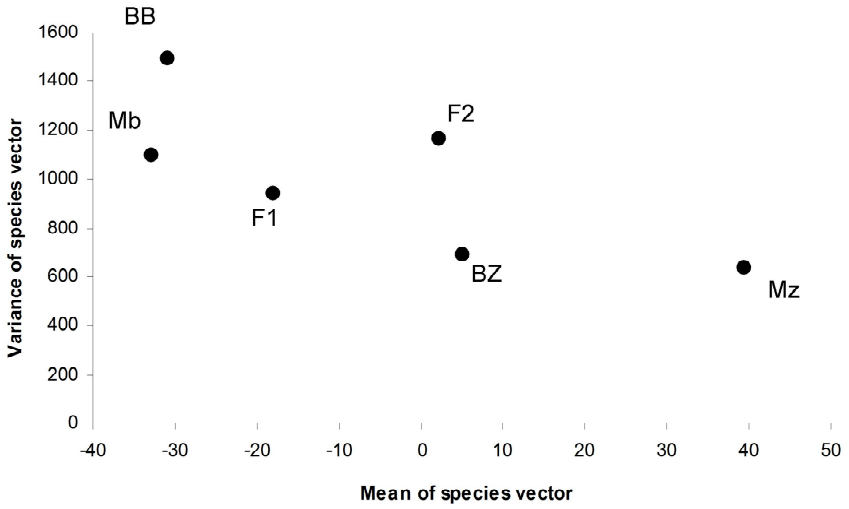

Figure 6.3. Mean vs. variance of the species vector plotted for the parental and hybrid generations.

Using the approach provided by Stelkens et al. (2009) our data show clear signs of transgressive segregation (Fig. 6.4, Table. 6.5). The range of the F2 phenotypes clearly exceeds the ranges of shape space on each PC of the combined parentals (Fig. 6.4). The amount of transgression found in the combined hybrid generations differed between 7 % (PC4) and 52.1 % (PC3). The total amount of transgression across all axes and adjusted for the variance explained by each axis was 23.8 %. The genetic p-distance between the two taxa calculated from 163 D-Loop sequences was 0.002 (SE 0.004).

The convex hull analysis confirmed our finding of transgressive segregation when adjusting for sample size (Fig. 6.5). All 95 % confidence intervals were extremely small and did not overlap. The two parental species are fairly similar in morphospace occupation. The F2 generation has a higher convex hull volume than the combined

volume of the parental species independent of the sample size. Thus the parental species occupy only a subset of the overall morphospace occupied by hybrids.

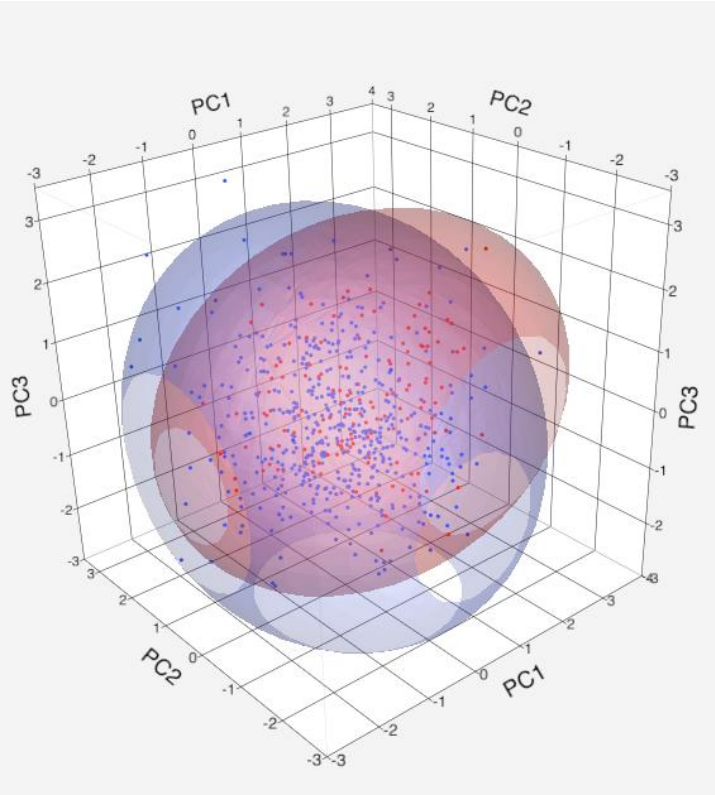

Figure 6.4. Transgressive segregation in body shape. PC1, PC2 and PC 3 for body shape were plotted for the parental (red) and hybrid generations (blue) with JMP v. 10. The spheres indicate 99 % normal contour ellipsoids. Note that only the first three PC axes are shown and therefore the total amount of transgressive segregation is not displayed here.

Discussion

In a previous study, we identified consistent differentiation in body shape in sympatric pairs of closely related species of the genus *Maylandia* (Husemann et al. in review). Here we used a common garden experiment and hybrid crosses to quantify the genetic and plastic contributions to body shape variation observed in natural populations.

We further used the cross to explore the mode of gene action and to test for transgressive segregation in the hybrids.

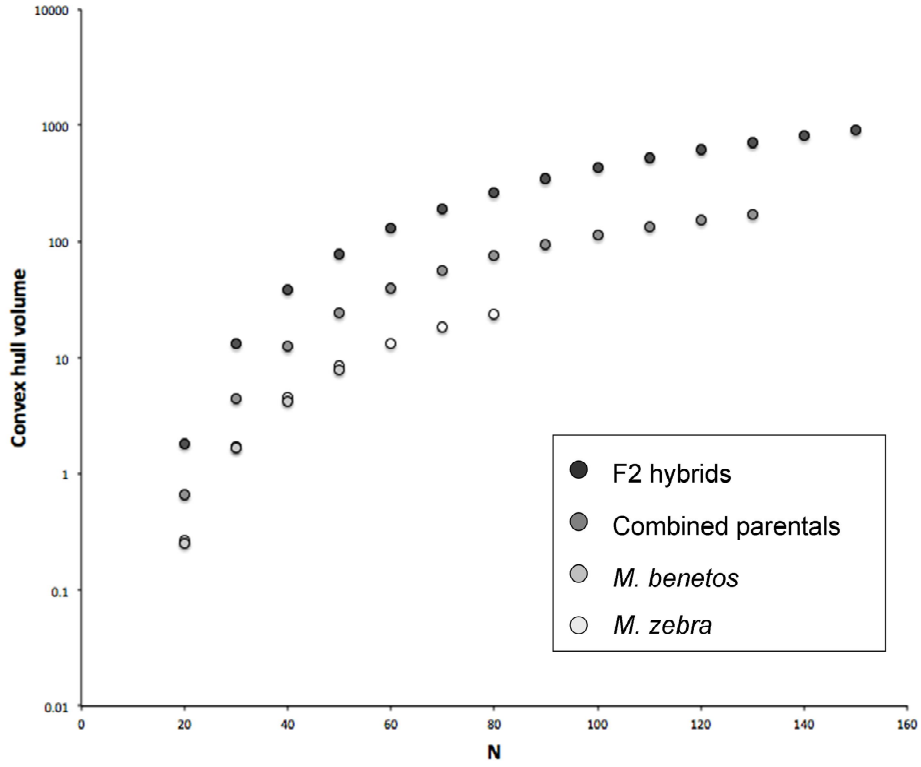

Figure 6.5. Comparison of morphospace occupation (Convex hull volume) in each of the parental species, both parentals combined and the F2 hybrid generation adjusted to different sample size using the Quickhull algorithm; 95% confidence intervals are too narrow to be visible (see online supplement Table for actual values).

Table 6.4. Amount of transgressive segregation found at each axis calculated according to Stelkens et al. (2009). Recorded is the amount of transgression in % found at each PC axis (TSPCi). The total amount for transgression across all axes and adjusted for the variance explained by the axis was 23.8 %.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
% explained	0.2091	0.1566	0.0953	0.0684	0.0604	0.0498	0.0445	0.0389	0.0360
parentals range	4.92237	4.2328	3.8416	5.7858	4.7983	4.5127	5.7604	4.5817	4.5934
total range	5.9971	6.1042	5.8426	6.1883	5.3630	6.7750	6.2585	6.2241	6.9664
TSPCi (%)	21.8	44.2	52.1	7	11.8	50.1	8.6	35.8	51.7

Genetic and Plastic Components of Body Shape

Our data clearly show that species differences in body shape have a genetic basis. Despite exhibiting strong plasticity in response to the rearing environment, the body shape differences observed between the two species are maintained after ~12 generations in a lab environment (Fig. 6.2). In addition, the cross data support a genetic component of body shape with the F1 and F2 generation being roughly intermediate between the two parental species (Fig. 6.3). This strong genetic component to body shape suggests that the species are adapted to different micro-niches in nature and that this adaptation is the result of habitat or community specific selective pressures.

Nonetheless, there is also a strong plastic component to body shape: both species have changed in body shape under laboratory conditions (Fig. 6.2). The changes followed a similar path, as both species got more slender with shallower bodies and thinner caudal peduncles. This might be a plastic response to smaller spaces, increased nutritional input, a structurally less diverse habitat and/or reduced predation pressure in the laboratory (Kerschbaumer et al. 2011).

The observed plasticity in body shape may play an adaptive role in the diversification of Lake Malawi cichlids. Phenotypic plasticity can facilitate rapid changes in morphology thereby allowing species to utilize novel environmental resources. This in turn might prevent competitive exclusion in complex and highly competitive communities like those observed in mbuna communities (Ghalambor et al. 2007, Ollson & Eklöv 2005). Furthermore, a plastic response can lead to heritable adaptive changes if selection favors a specific character state within the reaction norm across multiple generations (Via et al. 1995, see below).

A strong plastic component to body shape is not surprising since cichlids exhibit plasticity in a variety of morphological traits, such as trophic morphology (Meyer 1987, Muschick et al., 2011), body shape (Wimberger, 1992), mouth orientation, the size and orientation of fins, the thickness of the caudal peduncle (Kerschbaumer et al. 2011), and other morphological features (Magalhaes et al. 2009). Interestingly we find many similar traits to be plastic in our cross: body depth, orientation of the mouth and the thickness of the caudal peduncle show strong variation between environments (Fig. 6.2). These traits are important for feeding and swimming performance and therefore can be assumed to be under strong ecological selection (Langerhans & Reznik 2009). However, even in traits under strong selection high genetic variability can be maintained if strong genotype-by-environment (GxE) interactions occur (Greenfield et al. 2012). This in turn provides the genetic variation to quickly react to environmental change via plastic responses.

In our common garden experiment, we observed a low but significant GxE interaction (Table 2). GxE is the genetic variation in phenotypic plasticity and determines the reaction norm of an organism (Rodriguez 2012). A GxE interaction can provide the genetic variation required to promote the evolution of selectively advantageous phenotypic plasticity (Via & Lande 1985). Further, a plastic response can evolve into an adaptation in response to selection (Stearns 1989, Nussey et al. 2005). Thus strong GxE interactions are thought to help maintain phenotypic diversity even in the face of selection if the environmental conditions between isolated populations differ (Levene 1953, Johnson 2007, Greenfield et al. 2012). Our findings suggest that the observed GxE interactions and the resulting phenotypic plasticity may have contributed to the large phenotypic diversity observed in the mbuna and the larger East African cichlid radiation.

Analyses of Hybrid Crosses

Little is known about the genetic components of fish body shape and most of what is known comes from the stickleback model system. In sticklebacks the divergence in body shape between marine and freshwater lines appears to be determined by many genes (Schluter et al. 2004). A QTL analysis revealed that body shape in sticklebacks is determined by few genes with large effects in addition to multiple genes with smaller effects (Albert et al. 2007). In cichlids only the genetics of tropic morphology have been investigated (Albertson et al. 2003a, b, Albertson & Kocher 2005, Albertson & Kocher 2006, Fan et al. 2012). In these studies, one to eleven genetic factors have been determined to underlie shape differences of individual elements of trophic structures and pleiotropic effects appear to be a common feature in the genetic architecture of this trait set (Albertson et al. 2003a).

Despite our limited knowledge, it appears that body shape represents a composite trait set with a complicated genetic architecture that does not follow a simple additive model. Future quantitative genetic and developmental studies are required to understand the genetic basis of body shape in cichlids in specific. Such studies may shed light on the forces driving species divergence (Walker 2010).

Transgressive Segregation

Hybridization, especially with introduced species, is often considered a force leading to a decline of biodiversity because it disrupts species boundaries (e.g. Perry et al. 2002). However, the elevated genetic and phenotypic variance resulting from hybridization provides new variation that selection can act upon. In this way, hybridization can lead to the evolution of new adaptive phenotypes (e.g. Rieseberg et al.

2003b, Bell & Travis 2005, Seehausen 2004, Stelkens et al. 2009, Lucek et al. 2010, Genner & Turner 2012).

Tests for transgressive segregation in body shape show clearly that the hybrids occupy morphospace beyond what is found in the parentals indicating transgression (Table 6.5, Fig. 6.4, 6.5). The amount of transgression found in the cross was 23.8 % which is relatively high for a cross of two closely related species (Allender et al. 2003, Stelkens & Seehausen 2009, Stelkens et al. 2009, Husemann et al. in review).

The genetic distance between the two *Maylandia* species studied here was 0.002. This value is lower than any of the distances recorded in the study by Stelkens et al. (2009). However, the amount of transgression observed in our study exceeded estimates of transgression for F2 generations found by Stelkens et al. (2009) for taxa showing an order of magnitude higher genetic divergence. This suggests that hybridization can lead to high amounts of transgression even in very close relatives as commonly found in Lake Malawi cichlids.

Transgressive segregation is a common feature in many cichlid phenotypes. Albertson & Kocher (2005) for example have shown that the cichlid skull is susceptible to transgressive segregation, which is in line with our findings, though transgressive segregation in the jaw seems limited by the genetic architecture of this phenotype. Parsons et al. (2011) confirmed high degrees of transgressive segregation in head shape in a pair of Malawi rock-dwellers. Further, transgressive segregation has been detected for a coloration phenotype in Lake Malawi rock-dwellers (O'Quin et al. 2012). This high potential for transgressive segregation in different phenotypes suggests that hybridization can promote evolvability in East African cichlids and might be an important mechanism

in generating new diversity for selection to act on (Salzburger et al. 2002, Seehausen 2004, Bell & Travis 2005, Stelkens et al. 2009, Parsons et al. 2011, Magalhaes et al. 2012).

Conclusion

Our analysis of body shape in wild caught and laboratory reared specimens of two closely related species of cichlids revealed that species specific differences have a genetic basis. In addition body morphology has a plastic component and a small, yet significant GxE interaction. Thus body shape has the necessary genetic variation and plastic response needed to promote and maintain diversity. The mode of gene action of the species differences is complex, likely polygenic, and involves dominant and epistatic interactions. The potential for transgressive segregation is high supporting the possibility of an important role of hybridization in cichlid diversification.

CHAPTER SEVEN

Summary and Conclusions

Selection appears to be the driving force of phenotypic evolution in rock-dwelling cichlids. However, small population sizes of microendemics expose them to the effects of drift. Evidence for correlational selection suggests that specific combinations of reproductive and ecological traits might be favored, thus facilitating the divergence and coexistence of closely related species. Finally, phenotypic plasticity and transgressive segregation in hybrids are two mechanisms generating new phenotypic diversity, which may have contributed to the extraordinary diversification of this evolutionary model system.

APPENDICES

APPENDIX A

Table A.1. Microsatellites used in this study: given is their repeat pattern and motive, the range of allele sizes, the annealing temperature, the linkage group on which the locus is found, the primers used and the reference where primers were taken from.

ID	Bp repeat	Repeat pattern	Allele sizes	Annealing temp.	Genome location	Primer f	Primer r	reference
UNH2037	2	CA	126-236	58°C	LG 2	GGGATTCACTGGCACCTACT	ATGTGGTTCCCAGTGATGGT	Albertson et al. 2003
UNH2086	2	CA	134-282	59°C	LG 1	AACAGGCCGAGCAGAAAGT	CGATAGGCTGTTTTCCTGGAG	Albertson et al. 2003
UNH2139	2	GT	195-259	56°C	LG 12	GCAGTGCACATGCGACTTAT	ACAGCCAGCTACTGTGCAAC	Albertson et al. 2003
UNH2204	2	GT	116-194	56°C	LG 8	CACATCATGTCAATCAGACATCC	GGAGACGGTTCAAAGTCCTG	Albertson et al. 2003
UNH2190	2	GT	120-194	56°C	LG 10	GTTCGGCTGTGATGGTGATT	AGCGAGGACGGAGCTTTAAC	Albertson et al. 2003
UNH2169	2	GT	115-203	56°C	LG 11	CCAGTGGGTCCTCCTACAGA	CCCAGTGACTTTGAGGTGTG	Albertson et al. 2003
UNH2112	2	GT	121-227	56°C	LG 13	CTCGGTGGTCAGAATGAAGG	TTACAGCACTTCACGGTTGC	Albertson et al. 2003
UNH362	2	CA	110-188	56°C	LG 17	GAACAGCTTTCAGACGGAGG	ACTGAGGCCAGGTGAAGAAA	Kocher unpubl.
UNH2166	2	CA	119-233	56°C	LG 16	ACTGGCCCAAAACTGTCAAA	TGTGTGCCAAGGATAGCAAA	Albertson et al. 2003
UNH2065	2	CA	109-221	56°C	LG 19	CCGGGATGATTTTCTCACTG	CAGCACACGACAGGAGGTT	Albertson et al. 2003
UNH2152	2	GT	117-261	56°C	LG 3	TGACTGCTGCACATTATAACTCC	CAGCATGAAACTCACTGGAAA	Albertson et al. 2003
UNH2135	2	GT	105-243	56°C	LG 4	CCTGACAAAGCTGATTGTTCC	GTGAATGCTGAGGCAAGTCA	Albertson et al. 2003
UNH231	2	CA	150-282	56°C	LG 6	GCCTATTAGTCAAAGCGT	ATTTCTGCAAAAGTTTTCC	Lee & Kocher unpubl.

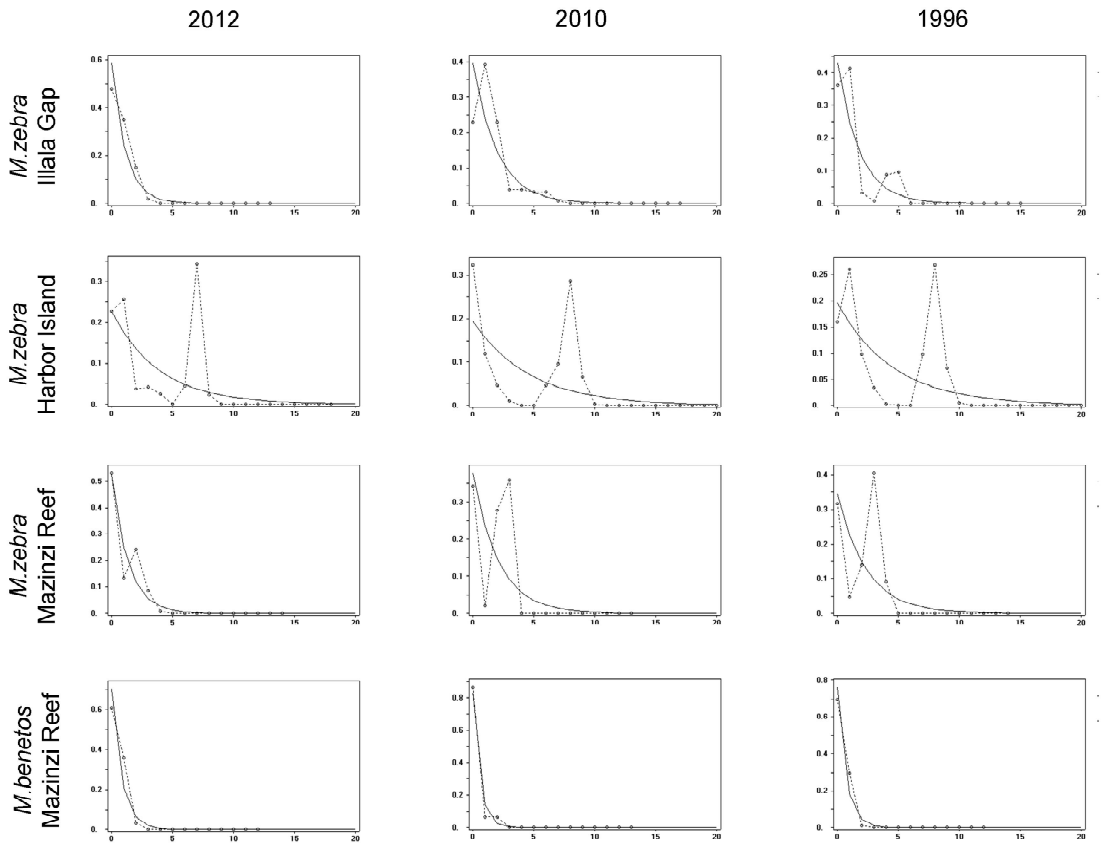

Figure A.1. Mismatch Distributions for D-Loop data testing for population demographic history generated with Arlequin

APPENDIX B

Table B.1. Laboratory and field comparison for melanophore counts and body shape to investigate the plastic and heritable components of the traits.

MANCOVA body shape	Hypothesis d.f.	Wilk's Lambda	F	p-value	η^2_p
Standard length	9.000	0.822	4.872	**<0.001**	0.178
Environment	9.000	0.415	31.671	**<0.001**	**0.585**
Sex	9.000	0.858	3.713	**<0.001**	0.142
Species	9.000	0.600	14.990	**<0.001**	**0.400**
Environment * Species	9.000	0.644	12.409	**<0.001**	**0.356**
Sex * Species	9.000	0.930	1.694	0.092	0.070
Environment * Sex	9.000	0.877	3.162	**0.001**	0.123

MANOVA melanophores	d.f	Sum of Squares	F	p-value
environment	1	690	1.287	0.260
Species	1	85993	160.347	**<0.001**
Environment*species	1	671	1.250	0.267
residuals	77	41294		

APPENDIX C

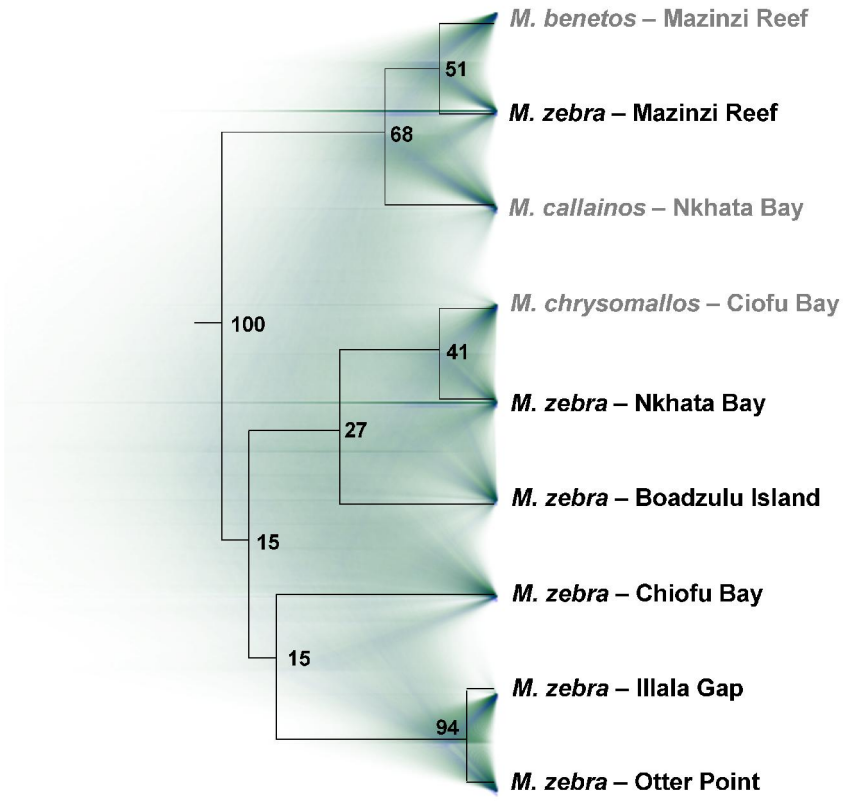

Figure C.1. Densitree visualization of all trees resulting from *BEAST analysis; the consensus tree and posterior probabilities are also shown; grey font represent non-barred phenotype, black is barred phenotype.

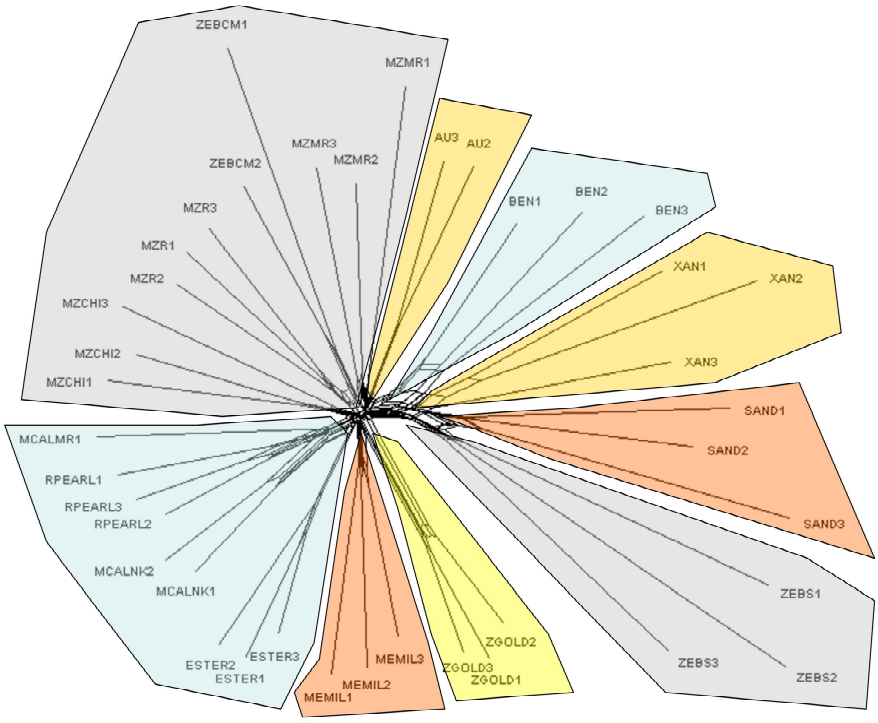

Figure C.2. Re-analyses of AFLP data from Allender *et al*. (2003). We re-analyzed part of the data provided by Allender *et al.* (2003). The original dataset contained 2,189 polymorphic AFLP sites sampled for a total of 58 individuals belonging to 25 taxa (including 6 outgroups). For our network analysis we only included members of the genus *Maylandia* which resulted in a reduced dataset of 40 individuals. We used the program Splitstree v. 4.11.3 (Huson & Bryant, 2006) to construct a neighbor-net using the equal angle method. The analysis indicates that different coloration phenotypes (grey – blue with black vertical bars, orange – blue with orange gular, blue – blue without bars, red – blue with red dorsal fin, yellow – golden body with light bars) have evolved independently from each other.

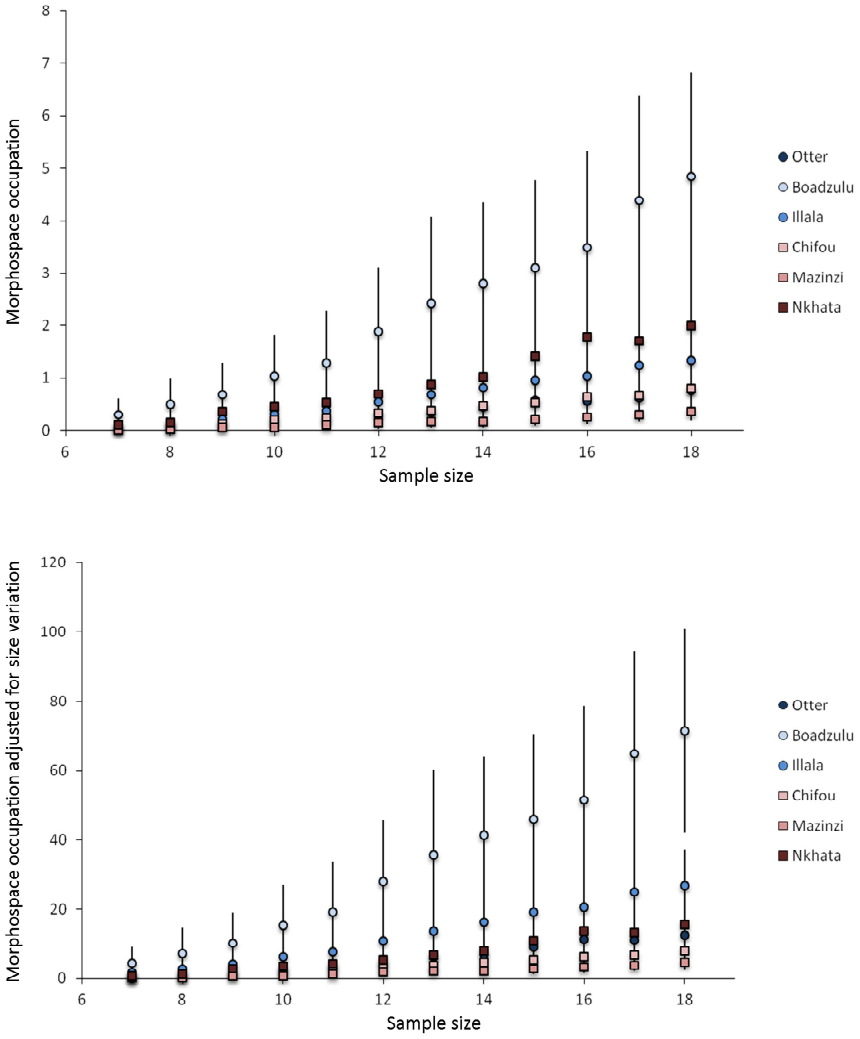

Figure C.3: Comparative analyses of morphospace across all sampled populations. To estimate morph space occupation for different populations of *M. zebra*, individual scores of the 9 principal component axis describing shape variation in 'sympatric' and 'allopatric' *M. zebra* were used to calculate the convex hull volume for each population using the Quickhull algorithm (Barber *et al.*, 1999). This analysis was performed both for the untreated PC scores and for size-corrected PC scores (size correction was conducted by means of a preparatory MANCOVA with the PC scores as dependent variables and size as a covariate; residuals were then used for further analysis). Due to different sample sizes among groups, we used a randomization procedure to calculate morphospace as described in Tobler & Carson (2010). Random distributions of morphospace occupation were generated using 100 iterations of randomly selecting (with replacing individuals after being drawn) a number of N specimens from the respective pool of individuals. Convex hull was calculated for each sample. Based on the 100 iterations means and confidence intervals were calculated for each group. Our analyses indicated that allopatric *M. zebra* populations tended to occupy a larger spectrum in morphospace (particularly when size-correcting the shape variables first). This pattern was mainly driven by the population from Boadzulu, which consistently had the highest morphospace occupation. The large overlap in morphospace occupation among populations, however, prevents us from drawing any conclusions; higher sample sizes – both in terms of populations and specimens within populations – are required to rigorously address this question.

REFERENCES

Abdallah AM, Barton DR. 2003. Environmental factors controlling the distributions of benthic invertebrates on rocky shores of Lake Malawi, Africa. J Great Lakes Res 29: 202-215

Adams DC. 2010. Parallel evolution of character displacement driven by competitive selection in terrestrial salamanders. BMC Evol Biol 10: 72.

Adams DC, Rohlf FJ, Slice DE. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. Ital J Zool 71: 5-16.

Aguilée R, Claessen D, Lambert A. 2013. Adaptive radiation driven by the interplay of eco-evolutionary and landscape dynamics. Evolution 67: 1291-1306.

Aguilée R, Lambert A, Claessen D. 2011. Ecological speciation in dynamic landscapes. J Evol Biol 24: 2663-2677.

Albert AYK, Sawaya S, Vines TH, Knecht AK, Miller CT, Summers BR, Balabhadra S, Kingsley DM, Schluter D. 2007. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. Evolution 62: 76-85.,

Albertson RC. 2008. Morphological divergence predicts habitat partitioning in a Lake Malawi cichlid species complex. Copeia 2008: 690-699

Albertson RC, Kocher TD. 2006. Genetic and developmental basis of cichlid trophic diversity. Heredity 97: 211-221

Albertson RC, Markert JA, Danley PD, Kocher TD. 1999. Phylogeny of a rapidly diverging clade: the cichlid fishes of Lake Malawi, East Africa. Proc Nat Acad Sci 96: 5107-5110.

Albertson RC, Kocher TD. 2005. Genetic architecture sets limits on transgressive segregation in hybrid cichlid fishes. Evolution 59: 686-690.

Albertson RC, Streelman JT, Kocher TD. 2003a, Genetic basis of adaptive shape differences in the cichlid head. J Hered 94: 291-301.

Albertson RC, Streelman JT, Kocher TD. 2003b. Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. Proc Nat Acad Sci 100: 5252-5257.

Allender CJ, Seehausen O, Knight ME, Turner GF, Madean N. 2003. Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. Proc Nat Acad Sci 100: 14074-14079

Allendorf FW, Luikart G. 2007. Conservation and the genetics of populations. Blackwell, Malden.

Allendorf FW. 1986. Genetic drift and the loss of alleles versus heterozygosity. Zool Biol 5: 181–190.

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. Nature Rev Genet 11: 697-710.

Alò D, Turner TF. 2004. Effects of habitat fragmentation on effective population size in the endangered Rio Grande Silvery Minnow. Conserv Biol 19: 1138-1148.

Anderson AN. 2008. Not enough niches: non-equilibrial processes promoting species coexistence in diverse ant communities. Austral Ecol. 33: 211-220.

Angeloni F, Ouborg NJ, Leimu R. 2011. Meta-analysis on the association of population size and life history with inbreeding depression in plants. Biol Conserv 144: 35-43.

Ardren WR, Kapuscinski AR. 2003. Demographic and genetic estimates of effective population size (N_e) reveals genetic compensation in steelhead trout. Mol Ecol 12: 35-49.

Armbruster WS, Edwards ME, Debevec EM. 1994. Floral character displacement generates assemblage structure of western Australian triggerplants (*Stylidium*). Ecology 75: 315-329.

Armstrong RA, McGehee R. 1980. Competitive exclusion. Am Nat 115: 151-170.

Arnegard ME, Markert JA, Danley PD, Stauffer Jr. JR, Ambali AJ, Kocher TD. 1999. Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. Proc R Soc Lond B 266: 199-130.

Arnegard ME. 2009. Ongoing ecological divergence in an emerging genomic model. Mol Ecol 18: 2926-2929.

Awiti AO. 2011. Biological diversity and resilience: Lessons from the recovery of cichlid species in Lake Victoria. Ecol Soc 16: 9.

Ayre DJ, Hughes TP. 2004. Climate change, genotypic diversity and gene flow in reef-building corals. Ecol Lett 7: 273-278.

Baird NA, Etter PD, Atwood TA, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLOS One 3, e3376.

Ball SJ, Adams M, Possingham HP, Keller MA. 2000. The genetic contribution of single male immigrants to small, inbred populations: a laboratory study using *Drosophila melanogaster*. Heredity 84: 677-684.

Barber CB, Dobkin DP, Huhdanpaa H. 1996. The Quickhull algorithm for convex hulls. ACM Trans Math Software 22: 469-483.

Barker JSF. 2011. Effective population size of natural populations of *Drosophila buzzatii*, with a comparative evaluation of nine methods of estimation. Mol Ecol 20: 4452-4471.

Barlow GW. 1972. The attitude of fish eye-lines in relation to body shape and to stripes and bars. Copeia 1: 4-12.

Barluenga M, Stölting KN, Salzburger W, Muschick M, Meyer A. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439: 719-723.

Barson NJ, Knight ME, Turner GF. 2007. The genetic architecture of male colour differences between a sympatric Lake Malawi cichlid species pair. J Evol Biol, 20, 45-53.

Barton NH. 1996. Natural selection and random genetic drift as causes of evolution on islands. Phil Trans Roy Soc Lond B 351: 785-795.

Beaumont M. 2000. Readme_2mod [Documentation file]. Available with the program at http://www.rubic.rdg.ac.uk/~mab/software.html

Beerli P, Palczewski, M 2010. Unified Framework to Evaluate Panmixia and Migration Direction Among Multiple Sampling Locations. Genetics 185: 313-326.

Bell MA, Travis MP. 2005. Hybridization, transgressive segregation, genetic covariation, and adaptive radiation. Trends Ecol Evol 20: 358-361.

Bergstrom CA. 2007. Morphological evidence of correlational selection and ecological segregation between dextral and sinistral forms in a polymorphic flatfish, *Platichtys stellatus*. J Evol Biol 20: 1104-1114.

Bijlsma R, Bundgaard J, Boerema AC. 2000. Does inbreeding affect the extinction risk of small populations?: predictions from *Drosophila*. J Evol Biol 13: 502-514.

Blais J, Plenderleith M, Rico C, Taylor MI, Seehausen O, van Oosterhout C, Turner GF. 2009. Assortative mating among Lake Malawi cichlid fish populations is not simply predictable from male nuptial coloration. BMC Evol Biol 9: 53.

Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26: 1372-1373.

Boughman JW. 2001. Divergent sexual selection enhances reproductive isolation in sticklebacks. Nature 41: 944 – 948

Bouton N, de Visser J, Barel CDN. 2002. Correlating head shape with ecological variables in rock-dwelling haplochromines (Teleostei: Cichlidae) from Lake Victoria. Biol J Linn Soc 76: 39-48.

Bouton N, Witte F, van Alphen JJM, Schenk A, Seehausen O. 1999. Local adaptations in populations of rock-dwelling haplochromines (Pisces: Cichlidae) from southern Lake Victoria. Proc R Soc Lond B 266: 355-360.

Bouzat, J. L., Paige, K. N. & Lewin, H. A. (1998). The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. American Naturalist 152, 1–6.

Breinholt JW, van Buren R, Kopp OR, Stephen CL. 2009. Population genetic structure of an endangered Utah endemic, *Astragalus ampullarioides* (Fabaceae). J Botany 96: 661-667.

Brodie ED III. 1992. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. Evolution 46: 1284-1298.

Burke JM, Arnold ML. 2001.Genetics and the fitness of hybrids. Ann Rev Genet 35: 31-52.

Carlsbeek R, Irschick DJ. 2007. The quick and the dead: correlational selection on morphology, performance, and habitat use in island lizards. Evolution 61: 2493-2503.

Carstens BC, Knowles LL. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: An example from *Melanoplus* grasshoppers. Syst Biol 56: 400-411.

Chaloupka M, Kamezaki N, Limpus C. 2008. Is climate change affecting the population dynamics of the endangered Pacific loggerhead sea turtle? J Exp Mar Biol Ecol 356: 136-143.

Charlesworth B. 2009. Effective population size and patterns of molecular evolution and variation. Nature Rev Genet 10: 195-205.

Chenoweth SF, Blows MW. 2008. Q_{ST} meets the G matrix: the dimensionality of adaptive divergence in multiple correlated quantitative traits. Evolution 62: 1437-1449.

Ciccotto PJ, Konings A, Stauffer JR Jr. 2011. Descriptions of five new species in the genus *Metriaclima* (Teleostei: Cichlidae) from Lake Malawi, Africa. Zootaxa 2738: 1-25.

Clement M, Posada D, Crandall K. 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657-1660.

Condé B, Géry J. 1999. Maylandia Meyer et Foerster, 1984, unnom générique disponible (Teleostei, Perciformes, Cichlidae). Rev Franc Aqua Herp 26: 21-22.

Coombs JA, Letcher BH, Nislow KH. 2008. CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs. Mol Ecol Res 8: 578-580.

Coombs JA, Letcher BH, Nislow KH. 2012. GONe: Software for estimating effective population size in species with generational overlap. Mol Ecol Res 12: 160-163.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-2014.

Couldrige VCK, Alexander GJ. 2002. Color patterns and species recognition in four closely related species of Lake Malawi cichlid. Behav Ecol 18: 59-64

Coyne JA, Orr HA. 2004. Speciation. Sinauer Associates, Inc. Sunderland, Massachusetts, USA.

Crawford NG. 2010. SMOGD: software for the measurement of genetic diversity. Mol Ecol Res 10: 556-557.

Crispo E, Chapman LJ. 2010. Temporal variation in population genetic structure of a riverine African cichlid fish. J Hered 101: 97-106

Dangles O, Carpio C, Barragan AR, Zeddam J-L, Silvain JF. 2008. Temperature as a key driver of ecological sorting among invasive pest species in the tropical Andes. Ecol Appl 18: 1795-1809.

Danley PD. 2001. The ecological, behavioral, and genetic factors influencing the diversification of Lake Malawi's Rock Dwelling Cichlids. University of New Hampshire, Durham

Danley PD. 2011. Aggression in closely related Malawi cichlids varies inversely with habitat complexity. Environ Biol Fish 92: 275-284.

Danley PD, Husemann M, Chetta J. 2012b. Acoustic diversity in Lake Malawi's rock-dwelling cichlids. Environ Biol Fish 93: 23-30.

Danley PD, Husemann M, Ding B, DiPietro LM, Beverly EJ, Peppe DJ. 2012a. The impact of the geologic history and paleoclimate on the diversification of East African cichlids. Int J Evol Biol 2012: 574851.

Danley PD, Kocher TD. 2001. Speciation in rapidly diverging systems: Lessons from Lake Malawi. Mol Ecol 10: 1075-1086

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

Darwin C. 1859. The origin of species by means of natural selection, or the preservation of favored races in the struggle for life. Modern Library, New York.

Davey JW, Blaxter ML. 2010. RADSeq: next-generation population genetics. Brief Funct Genomics 9: 416-423.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Rev Genet 12: 499-510.

De Pristo MA, Banks E, Poplin RE, Garimella KV, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Garbiel SB, Sltshuler D, Daly MJ. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genet 43: 491-498.

Dean MD, Ballard JW. 2004. Linking phylogenetics with population genetics to reconstruct the geographic origin of a species. Mol Phyl Evol 32: 998–1009.

Derryberry EP, Seddon N, Claramunt S, Tobias JA, Baker A, Aleixo A, Brumfield RT. 2012. Correlated evolution of beak morphology and song in the neotropical woodcreeper radiation. Evolution 66: 2784-2797.

Deutsch JC. 1997. Colour diversification in Malawi cichlids: evidence for adaptation, reinforcement or sexual selection? Biol J Linn Soc 62: 1-14.

Dieckmann U, Doebeli M. 1999. On the origin of species by sympatric speciation. Nature 400: 354-357.

Dijkstra PD, Seehausen O, Groothuis TGG. 2005. Direct male-male competition can facilitate invasion of new colour types in Lake Victoria cichlids. Behav Ecol Sociobiol 58: 136-143.

Dijkstra PD, Seehausen O, Pierotti MER, Groothuis TGG. 2006. Male-male competition and speciation: aggression bias towards differently coloured rivals varies between stages of speciation in a Lake Victoria cichlid species complex. J Evol Biol 20: 496-502.

Dillon N, Austin AD, Bartowsky E. 1996. Comparison of preservation techniques for DNA extraction from hymenopterous insects. Insect Mol Biol 5: 21-24.

Ding B, Curole J, Husemann M, Danley PD (in review) Habitat complexity predicts community diversity of rock-dwelling cichlids in Lake Malawi, East Africa. Evol Ecol.

DiRienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. 1994. Proc Nat Acad Sci 91: 3166-3170.

Dobzhanski T. 1970. Genetics and the evolutionary process. Columbia University Press, New York and London

Dominey WJ. 1984. Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian *Drosophila*. In Echelle AA and I Kornfield (eds.), Evolution of fish species flocks. University of Maine at Orono Press, Orono, pp. 231-250

Drees C, Matern A, Rasplus J-Y, Terlutter H, Assmann T, Weber F. 2008. Microsatellites and allozymes as the genetic memory of habitat fragmentation and defragmentation in populations of the ground beetle *Carabus auronitens* (Col., Carabidae). J Biogeogr 35: 1937–1949.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol 22:1185–1192

Drummond AJ, Suchard MA, Xie D & Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29: 1969-1973.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C. et al. 2011. Geneious v5.4, Available at http://www.geneious.com/.

Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Res 4: 359-361.

Eckert CG, Samis KE, Lougheed SC. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Mol Ecol 17: 1170–1188.

Edelaar P, Burraco P, Gomez-Mestre I. 2011. Comparisons between Q_{ST} and F_{ST} – how wrong have we been? Mol Ecol 20: 4830-4839.

Ekblom R, Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107: 1-16.

Ellegren H. 2004. Microsatellites: Simple sequences with complex evolution. Nature Rev Genet 4: 435-445.

Ellstrand NC, Elam DR. 1993. Genetic consequences of small population size: implications for plant conservation. Ann Rev Ecol Syst 24: 217-242.

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

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10: 564-567.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Fan S, Elmer KR, Meyer A. 2011. Genomics of adaptation and speciationin cichlid fishes: recent advances and analyses in African and Neotropical lineages. Phil Trans Roy Soc B 367: 385-394.

Ficetola GF, Padoa-Schioppa E, Wang J, Garner TWJ. 2009. Polygyny, census and effective population size in the threatened frog, *Rana latastei*. Anim Conserv 13: 1-8.

Ficetola GF, Bonin A, Miaud C. 2008. Population genetics reveals origin and numbers of founders in a biological invasion. Mol Ecol 17: 773-782.

Fickel J, Bubliy OA, Stache A, Noventa T, Jirsa A, Heurich M. 2012. Crossing the border? Structure of the red deer (*Cervus elaphus*) population from the Bavarian-Bohemian forest ecosystem. Mamm Biol 77: 211-220.

Finger A, Kettle CJ, Kaiser-Bunbury CN, Valentin T, Mougal J, Ghazoul J. 2012. Forest fragmentation genetics in a formerly widespread island endemic tree: *Vateriopsis seychellarum* (Dipterocarpaceae). Mol Ecol 21: 2369–2382.

Finger A, Kettle CJ, Kaiser-Bunbury CN, Valentin T, Doudee D, Matatiken D, Ghazoul J. 2011. Back from the brink: potential for genetic rescue in a critically endangered tree. Mol Ecol 20: 3773–3784.

Fisher RA, 1930. The Genetical Theory of Natural Selection.Clarendon, Oxford.

Frankham R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genet Res 66: 95-107.

Frankham R. 2005. Genetics and extinction. Biol Conserv 126: 131-140.

Frankham R, Ballou JD, Briscoe DA. 2004. A primer of population genetics. Cambridge University Press, Cambridge.

Franklin IF, Frankham R. 1998. How large must populations be to retain evolutionary potential? Anim Conserv 1: 69-73.

Fryer G, Iles TD. 1972. The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution. Oliver and Boyd, Edinburgh.

Gassert F, Schulter U, Husemann M, Ulrich W, Rödder D, Hochkirch A, Engel E, Meyer J, Habel JC. 2013. From southern refugia to the northern range margin: genetic population structure of common wall lizard, *Podarcis muralis*. J Biogeogr 40: 1475-1489.

Gause GF. 1934. The struggle for existence. Williams & Wilkins, Baltimore.

Gavrilets S. 2004. Fitness landscapes and the origin of species. Princeton University Press, Princeton, NJ.

Gavrilets S, Losos JB. 2009. Adaptive radiation: Contrasting theory with data. Science 323: 732-737

Genner MJ, Botha A, Turner GF. 2006. Translocations of rocky habitat cichlid fishes to Nkhata Bay, Lake Malawi. J Fish Biol 69: 622-628.

Genner MJ, Knight ME, Haesler MP, Turner GF. 2010. Establishment and expansion of Lake Malawi rock fish populations after a dramatic late Pleistocene lake level rise. Mol Ecol 19: 170-182.

Genner MJ, Turner GF. 2005. The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. Fish Fish 6: 1-34

Genner MJ, Turner GF. 2012. Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. Mol Biol Evol 29: 195-206.

Genner MJ, Turner GF, Hawkins SJ. 1999. Resource control by territorial male cichlid fish in Lake Malawi. J Anim Ecol 68: 522-529.

Genner MJ, Cleary DFR, Knight ME, Michel E, Seehausen O, Turner GF. 2004. How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? J Biogeogr 31: 93–102.

Ghalambor CK, McKay JK, Carroll SP, Reznik DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and potential for contemporary adaptation in new environments. Funct Ecol 21: 394-407.

Gilbert MTP, Moore W, Melchior L, Worobey M. 2007. DNA extraction from dry museum beetles without conferring external morphological damage. PLoS ONE 2: e272.

Gomaa NH, Montesinos-Navarro A, Alonso-Blanco C, Pico FX. 2011. Temporal variation in genetic diversity and effective population size of Mediterranean and subalpine *Arabidopsis thaliana* populations. Mol Ecol 20: 3540-3554.

Gompert Z, Forister ML, Fordyce JA, Nice CC, Williamson RJ, Buerkle CA. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of *Lycaeides* butteflies. Mol Ecol 19: 2455-2473.

Gophen M, Ochumba PBO, Pollingher U, Kaufman LS. 1993. Nile Perch (*Lates niloticus*) invasion in Lake Victoria (East Africa). Verh Internat Limnol 25: 856-859.

Goudet J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. J Hered 86: 485-486.

Grant PR. 1981. Speciation and the adaptive radiation of Darwin's finches: The complex diversity of Darwin's finches may provide a key to the mystery of how intraspecific variation is transformed into interspecific variation. Am Sci 9: 653-663.

Grant PR, Grant BR. 2006. Evolution of character displacement in Darwin's finches. Science 313: 224-226.

Greenfield MD, Danka RG, Gleason JM, Harris BR, Zhou Y. 2012. Genotype x environment interaction, environmental heterogeneity and the lek paradox. J Evol Biol 25: 601-613.

Griffith AM, Koizumi I, Bright D, Stevens JR. 2009. A case of isolation by distance and short-term temporal stability of population structure in brown trout (*Salmo trutta*) within the River Dart, southwest England. Evol Appl 2: 537-554.

Groombridge JJ, Jones CG, Bruford MW, Nichols RA. 2000. 'Ghost' alleles of the Mauritius kestrel. Nature 403, 616.

Habel JC, Husemann M, Finger A, Danley PD, Zachos FE. in press. The relevance of time series in molecular ecology and conservation biology. Biol Rev.

Habel JC, Husemann M, Schmitt T, Zachos FE, Honnen A-C, Peterson B, Parmakelis, Stathi I. 2012. Microallopatry caused strong diversification in *Buthus* scorpions (Scorpiones: Buthidae) in the Atlas Mountains (NW Africa). PLoS One 7: e29403.

Habel JC, Zachos FE. 2012. Habitat fragmentation versus fragmented habitats. Biodiv Conserv 21: 2987-2990.

Habel JC, Zachos FE, Finger A, Meyer M, Louy D, Assmann T, Schmitt T. 2009. Unprecedented long-term genetic monomorphism in an endangered relict butterfly species. Conserv Genet 10: 1659– 1665.

Habel JC, Finger A, Schmitt T, Nève G. 2011. Changing over time: 15 years of isolation influence the population genetic structure of the endangered butterfly *Lycaena helle*. J Zool Syst Evol Res 49: 25-31.

Habel JC, Mulwa RK, Gassert F, Rödder D, Meyer J, Lens L. 2013. Genetic signature of diverging habitat features in the Mountain White-eye *Zosterops poliogaster*. Heredity.

Habel JC, Reuter M, Drees C, Pfaender J. 2012. Does isolation affect phenotypic variability and fluctuating asymmetry in the endangered Red Apollo? J Insect Conserv 16: 571-579.

Hangartner S, Laurila A, Räsänen K. 2011. Adaptive divergence in moor frog (*Rana arvalis*) populations along an acidification gradient: inferences from $Q_{ST} - F_{ST}$ correlations. Evolution 66: 867-881.

Hanski I. 1999. Metapopulation Ecology. Oxford University Press.

Hansson B, Westerberg L. 2002. On the correlation between heterozygosity and fitness in natural populations. Mol Ecol 11: 2467–2474.

Harpending HC. 1994. Signature of ancient population growth in a low-resolution mitochondrial mismatch distribution. Human Biology 66: 591-600.

Harper LG, Maclean N, Goulson D. 2006. Analysis of museum specimens suggests extreme genetic drift in the adonis blue butterfly (*Polyommatus bellargus*). Biol J Linn Soc 88: 447–452.

Hartl DL, Clark AG. 2007. Principles of population genetics. Sinauer Associates, Sunderland, Massachusetts.

Hauser L, Adock GJ, Smith PJ, Bernal Ramirez JH, Carvalho GR. 2002. Loss of microsatellite diversity and low effective population in an overexploited population of New Zealand Snapper (*Pagrus auratus*). Proc Nat Acad Sci 99: 11742-11747.

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

Hein J, Schierup MH, Wiuf C. 2005. Gene Genealogies, Variation and Evolution: A Primer in Coalescent Theory. Oxford University Press, USA.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. Mol Biol Evol 27: 570-580.

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

Herler J, Kerschbaumer M, Mitteroecker P, Postl L, Sturmbauer C. 2010. Sexual dimorphism and population divergence in the Lake Tanganyika cichlid fish genus *Tropheus*. Front Zool 7: 4.

Hert E. 1989. The function of egg-spots in an African mouth-brooding cichlid fish. Anim Behav 37: 726-732.

Higgins SN, Kling HJ, Hecky RE, Taylor WD, Bootsma HA. 2003. The community composition, distribution, and nutrient status of epilithic periphyton at five rocky littoral zones sites in Lake Malawi, Africa. J Great Lakes Res 29: 181-189

Hoarau G, Boon E, Jongma DN, Ferber S, Palsson J, Van der Veer HW, Rijnsdorp AD, Stam WT, Olsen JL. 2005. Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice (*Pleuronectes platessa* L.). Proc Biol Sci 272: 497-503.

Hoeck PEA, Bollmer JL, Parker PG, Keller LF. 2010. Differentiation with drift: a spatio-temporal genetic analysis of Galápagos mockingbird populations. Phil Trans Roy Soc B 365: 1127-1138.

Hoffman EA, Schueler FW, Blouin MS. 2004. Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern Leopard frog. Evolution 58: 2536-2545.

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. 2002. Ancient DNA. Nature Rev Genet 2: 353-359.

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G. 2011. Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Mol Ecol Res 11: 117-122.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet 6: e1000862.

Holzberg S. 1978. A field and laboratory study of the behaviour and ecology of *Pseudotropheus zebra* (Boulenger), an endemic cichlid of Lake Malawi (Pisces: Cichlidae). J Zool Syst Evol Res 16: 171-187.

Honnay O, Jacquemyn H. 2007. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. Conserv Biol 21: 823–831.

Honnay O, Jacquemyn H, Van Looy K, Vandepitte K, Breyne P. 2009. Temporal and spatial genetic variation in a metapopulation of the annual *Erysimum cheiranthoides* on stony river banks. J Ecol 97: 131–141.

Hoy RR, Hoikkala A, Kaneshiro K. 1988. Hawaiian courtship songs: Evolutionary innovation in communication signals of *Drosophila*. Science 240: 217-219.

Husemann M, Ray JW, King RS, Hooser EA, Danley PD. 2012. Comparative biogeography reveals differences in population genetic structure of five species of stream fish. Biol J Linn Soc 107: 867-885.

Husemann M, Schmitt T, Stathi I, Habel JC. 2012. Evolution and radiation in the scorpion *Buthus elmoutaouakili* Lourenco and Qi 2006 (Scorpiones: Buthidae) at the foothills of the Atlas Mountains. J Hered 103: 221-229.

Husemann M, Tobler M, McCauley C, Ding B, Danley PD. In review. The correlation of reproductive and ecological traits may facilitate the coexistence of congeners in Lake Malawi's rock-dwelling cichlids. J Evol Biol.

Husemann M, Namkung S, Habel JC, Danley PD, Hochkirch A. 2012. Phylogenetic analyses of band-winged grasshoppers (Orthoptera, Acrididae, Oedipodinae) reveal convergence of wing-morphology. Zool Scr 41: 515-526.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23: 254–267.

Ingram T, Svanbäck R, Kraft NJB, Kratina P, Southcott L, Schluter D. 2012. Intragulid predation drives evolutionary niche shift in threespine stickleback. Evolution 66: 1819-1832.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801-1806

Jensen JL, Bohonak AJ, Kelley ST. 2005. Isolation by distance, web service. BMC Genet 6: 13.

Johnson MTJ. 2007. Genotype-by-enviroment interactions lead to variable selection on life-history strategy in Common Evening Primrose (*Oenothera biennis*). J Evol Biol 20: 190-200.

Johnstone DL, O'Connell MF, Palstra FP, Ruzzante DE. 2012. Mature male pall contribution to the effective size of an anadromous Atlantic salmon (*Salmo salar*) population over 30 years. Mol Ecol 22: 2394-2407.

Jorde PE, Ryman N. 1995. Temporal allele frequency change and estimation of effective size in populations with overlapping generations. Genetics 139: 1077-1090.

Jorde PE, Ryman N. 2007. Unbiased estimator for genetic drift and effective population size. Genetics 177: 927-935

Joyce DA, Lunt DH, Bills R, Turner GF, Katongo C, Duftner N, Sturmbauer C, Seehausen O. 2005. An extinct cichlid fish radiation emerged in an extinct Pleistocene lake. Nature 435: 90-95.

Joyce DA, Lunt DH, Genner MJ, Turner GF, Bills R, Seehausen O. 2011. Repeated colonization and hybridization in Lake Malawi cichlids. Curr Biol 21: 108-109.

Kadlec T, Vrba P, Kepka P, Schmitt T, Konvička M. 2010. Tracking the decline of once-common butterfly: delayed oviposition, demography and population genetics in the Hermit, *Chazara briseis*. Anim Conserv 13: 172-183.

Kaplan I, Denno RF. 2007. Interspecific interactions in phytophagous insects revisited: a quantitative assessment of competition theory. Ecol Lett 10: 977-994.

Kass RE, Raftery AE. 1995. Bayes factors. J Am Stat Assoc 90: 773-795.

Kaufman LS. 1992. Catastrophic change in species-rich freshwater ecosystems, the lessons of Lake Victoria. Bioscience 42:846.

Keller I, Largiadèr CR. 2003. Recent habitat fragmentation caused by major roads leads to reduction of gene flow and loss of genetic variability in ground beetles. Proc Roy Soc Lond B 270: 417–423.

Keller I, Excoffier L, Largiadèr CR. 2005. Estimation of effective population size and detection of a recent population decline coinciding with habitat fragmentation in a ground beetle. J Evol Biol 18: 90–100.

Keller LF, Waller DM. 2002. Inbreeding effects in wild populations. Trends Ecol Evol 17: 230–241.

Kerschbaumer M, Sturmbauer C. 2011. The utility of geometric morphometrics to elucidate pathways in cichlid fish evolution. Int J Evol Biol 2011: 290245.

Kerschbaumer M, Postl L, Koch M, Wiedl T, Sturmbauer C. 2011. Morphological distinctness despite large-scale phenotypic plasticity – analysis of wild and pond-bred juveniles of allopatric populations of *Tropheus moorii*. Naturwissenschaften 98: 125-134.

Kidd MR, Danley PD, Kocher TD. 2006. A direct assay of female choice in cichlids: all eggs in one basket. J Fish Biol 68: 373-384.

Kimura M, Crow JF. 1964. The number of alleles that can be maintained in a finite population. Genetics 49: 725-738.

Knight ME, Turner GF. 2004. Laboratory mating trails indicate incipient speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra* from Lake Malawi. Proc R Soc Lond B 271: 675-680

Knight ME, Turner GF, Rico C, van Oppen MJH, Hewitt GM. 1998. Microsatellite analysis on captive Lake Malawi cichlids supports reproductive isolation by direct mate choice. Mol Ecol 7: 1606-1610.

Knowles LL. 2001. Genealogical portraits of speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of the Rocky Mountains. Proc Roy Soc Lond B 268: 319-324.

Koblmüller S, Duftner N, Sefc KM, Aibara M, Stipacek M, Blanc M, Egger B, Sturmbauer C. 2007. Reticulate phylogeny of gastropod-shell-breeding cichlids from Lake Tanganyika – the result of repeated introgressive hybridization. BMC Evol Biol 7: 7.

Koblmüller S, Salzburger W, Obermüller B, Eigner E, Sturmbauer C, Sefc KM. 2011. Separated by sand, fused by dropping water: habitat barriers and fluctuating water levels steer the evolution of rock-dwelling cichlid populations in Lake Tanganyika. Mol Ecol 20: 2272-2290.

Koblmüller S, Schliewen UK, Duftner N, Sefc KM, Katongo C, Sturmbauer C. 2008. Age and spread of the haplochromine cichlid fishes in Africa. Mol Phyl Evol 49: 153-169.

Koch M, Wilson AJ, Kerschbaumer M, Wiedl T, Sturmbauer C. 2012. Additive genetic variance of quantitative traits in natural and pond-bred populations of the Lake Tanganyika cichlid *Tropheus moorii*. Hydrobiologia 682: 131-141.

Kocher TD. 2004. Adaptive evolution and explosive speciation: the cichlid model. Nature Rev Genet 5: 288-298.

Konvicka M, Benes J, Schmitt T. 2010. Ecological limits vis a vis changing climate: relic Erebia butterflies in insular Sudetan mountains. In: Habel JC, Assmann T. (eds.). Relict species – phylogeography and conservation of relict species. Springer, Heidelberg: 341-355.

Koplin JR, Hoffmann RS. 1968. Habitat overlap and competitive exclusion in voles (*Microtus*). Am Midl Nat 80: 494-507.

Kornfield I, Smith PF. 2000. African cichlid fishes: model systems for evolutionary biology. Ann Rev Ecol Syst 31: 163-182.

Kristensen TN, Loeschke V, Hoffmann AA. 2008. Linking inbreeding effects in captive populations with fitness in the wild: Release of replicated *Drosophila melanogaster* lines under different temperatures. Conserv Biol 22: 189–199.

Lande R. 1976. Natural Selection and random genetic drift in phenotypic evolution. Evolution 30: 314-334.

Lande R. 1992. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. Evolution 46: 381–389.

Lande R, Schemske DW. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39: 24-40.

Langella O. 1999. Populations 1.2.31. available at http://bioinformatics.org/~tryphon/populations/

Langerhans RB, Reznick DN. 2009. Ecology and Evolution of swimming performance in fishes: predicting evolution with biomechanics. Pages 200-248 in: Domenici P, Kapoor BG, eds., Fish Locomotion: an etho-ecological perspective. Science Publishers, Enfield.

Langerhans RB. 2009. Trade-off between steady and unsteady swimming underlies predator-driven divergence in Gambusia affinis. J Evol Biol 22: 1057-1975.

Leimu R, Mutikainen P. 2005. Population history, mating system, and fitness variation in a perennial herb with a fragmented distribution. Conserv Biol 19: 349–356.

Leimu R, Mutikainen P, Koricheva J, Fischer M. 2006. How general are positive relationships between plant population size, fitness and genetic variation? J Ecol 94: 942–952.

Leinonen T, Cano JM, Mäkinen H, Merilä J. 2006. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. J Evol Biol 19: 1803-1812.

Leinonen T, McCairns RJS, O'Hara RB, Merilä J. 2013. $Q_{ST} - F_{ST}$ comparisons : evolutionary and ecological insights from genomic heterogeneity. Nature Rev Genet 14: 179-190.

Leinonen T, O'Hara RB, Cano JM, Merilä J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. J Evol Biol 21: 1-17.

Lermen D, Blömeke B, Browne R, Clarke A, Dyce PW, Fixemer T, Fuhr GR, Holt WV, Jewgenow K, Lloyd RE, Lötters S, Paulus M, Reid GM, Rapoport DH, Rawson D, Ringleb J, Ryder OA, Spörl G, Schmitt T, Veith M, Müller P. 2009. Cryobanking of viable biomaterials - implementation of new strategies for conservation purposes. Mol Ecol 18: 1030–1033.

Levene H. 1953. Genetic equilibrium when more than one ecological niche is available. Am Nat 87: 331–333.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Liem KF. 1973. Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Syst Zool 22: 425-441.

Liem KF. 1980. Adaptive significance of intra- and interspecific differences in the feeding repertoires of cichlid fishes. Am Zool 20: 25-31.

Loh Y-HE, Katz LS, Mims MC, Kocher TD, Yi SV, Streelman JT. 2008. Comparative analysis reveals signatures of differentiation amid genomic polymorphism in Lake Malawi cichlids. Genome Biol 9: R113.

Losos JB. 1990. A phylogenetic analysis of character displacement in Caribbean *Anolis* lizards. Evolution 44: 558-569.

Losos JB, Glor RE, Kolbe JJ, Nicholson K. 2006. Adaptation, speciation, and convergence: a hierarchical analysis of adaptive radiation in Caribbean *Anolis* lizards. Ann Missouri Bot Garden 93: 24-33.

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

Lucek K, Roy D, Bezault E, Sivasundar A, Seehausen O. 2010. Hybridization between distant lineages increases adaptive variation during a biological invasion: stickleback in Switzerland. Mol Ecol 19: 3995-4011

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conserv Genet 11: 355-373.

Luikart G, England PR, Tallmon D, Jordan S, Taberle P. 2003. The power and promise of population genomics: from genotyping to genome typing. Nature Rev Genet 4: 981-994.

Lynch M, Lande R. 1997. The critical effective size for a genetically secure population. Anim Conserv 1: 70-72.

Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, Massachusetts, USA.

Lynch M, Burger R, Butcher D, Gabriel W. 1993. The mutational meltdown in asexual populations. J Hered 84: 339-344.

Maan ME, Seehausen O, Van Alphen JJM. 2010. Female preferences and male coloration covary with water transparency in a Lake Victoria cichlid fish. Biol J Linn Soc 99: 398-406.

Magalhaes IS, Mwaiko S, Schneider MV, Seehausen O. 2009. Divergent selection and phenotypic plasticity during incipient speciation in Lake Victoria cichlid fish. J Evol Biol 22: 260-274.

Magalhaes IS, Lundsgaard-Hansen B, Mwaiko S, Seehausen O. 2012. Evolutionary divergence in replicate pairs of ecotypes of Lake Victoria cichlid fish. Evol Ecol Res 14: 381-401.

Manel S, Holderegger R. 2013. Ten years of landscape genetics. Trends Ecol Evol.

Marek PE, Bond JE. 2006. Phylogenetic systematics of the colorful, cyanide-producing millipedes of Appalachia (Polydesmida, Xystodesmidae, Apheloriini) using a total evidence Bayesian Approach. Mol Phyl Evol 41: 704-729.

Markert JA, Arnegard ME, Danley PD, Kocher TD. 1999. Biogeography and population genetics of the Lake Malawi cichlid Melanochromis auratus: habitat transience, philopatry and speciation. Mol Ecol 8: 1013-1026.

Martin CH, Genner MJ. 2009. High niche overlap between two successfully coexisting pairs of Lake Malawi cichlid fishes. Can J Fish Aquat Sci 66: 579-588.

McCairns RJS Bernatchez L. 2012. Plasticity and heritability of morphological variation within and between parapatric stickleback demes. J Evol Biol 25: 1097-1112.

McDonald KS III, Yampolsky L, Duffy JE. 2005. Molecular and morphological evolution of the amphipod radiation of Lake Baikal. Mol Phyl Evol 35: 323-343.

McGlothlin JW, Parker PG, Nolan VJr, Ketterson ED. 2005. Correlational selection leads to genetic interaction of body size and an attractive plumage trait in dark-eyed juncos. Evolution 59: 658-671.

McKay JK, Latta G. 2002. Adaptive population divergence: markers, QTL and traits. Trends Ecol Evol 17: 285-291.

McKaye KR, Louda SM, Stauffer JR Jr. 1990. Bower size and male reproductive success in a cichlid fish lek. Am Nat 135: 597-613.

Meglécz E, Petenian F, Danchin E, Coer D'Acier A, Rasplus J-Y, Faure E. 2004. High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydrias aurinia*. Mol Ecol 13: 1693–1700.

Merilä J, Crnokrak P. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. J Evol Biol 14: 892–903.

Meyer A. 1987. Phenotypic plasticity and heterochrony in *Cichlasoma managuense* (Pisces, Cichlidae) and their implications for speciation in cichlid fishes. Evolution 41: 1357-1369.

Meyer MK, Foerster W. 1984. Un nouveau Pseudotropheus du lac Malawi avec des remarques sur le complexe *Pseudotropheus-Melanochromis* (Pisces, Perciformes, Cichlidae). Rev Franc Aqua Herp 10: 107-112.

Miller LM, Kapuscinski AR. 1997. Historical analysis of genetic variation reveals low effective population size in a Northern Pike (*Esox lucius*) population. Genetics 147: 1249-1258.

Miller CR, Waits LP. 2003. The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): Implications for conservation. Proc Nat Acad Sci USA 100: 4334–4339.

Morin PA, Luikert G, Wayne RK. 2004. SNPs in ecology, evolution and conservation. Trends Ecol Evol 19: 208-216.

Morris AB, Baucom RS, Cruzan MB. 2002. Stratified analysis of the soil seed bank in the cedar glade endemic *Astragalus bibullatus*: evidence for historical changes in genetic structure. Am J Bot 89: 29–36.

Mundy NI, Unitt P, Woodruff DS. 1997. Skin from feet of museum specimens as a non-destructive source of DNA for Avian genotyping. Auk 114: 126-129.

Muschick M, Barluenga M, Salzburger W, Meyer A. 2011. Adaptive phenotypic plasticity in the Midas cichlid fish pharyngeal jaw and its relevance in adaptive radiation. BMC Evol Biol 11: 116.

Neff BD, Fu P, Gross MR. 1999. Microsatellite evolution in sunfish. Can J Fish Aquat Sci 56: 1198-1205.

Nevado B, Mautner S, Sturmbauer C, Verheyen E. 2013. Water-level fluctuations and metapopulation dynamics as drivers of genetic diversity in populations of three Tanganyikan cichlid fish species. Mol Ecol.

Nielsen EE, Hansen MM. 2008. Waking the dead: the value of population genetic analyses of historical samples. Fish Fish 9: 450–461.

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D. 2009. Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18: 3128-3150.

Nosil P. 2008. Speciation with gene flow could be common. Mol Ecol 17: 2103-2106.

Nussey DH, Postma E, Gienapp P, Visser ME. 2005. Selection on heritable phenotypic plasticity in a wild bird population. Science 310: 304-306.

O'Hara RB, Merilä J. 2005. Bias and precision in Q(ST) estimates: problems and some solutions. Genetics 171: 1331-1339.

O'Hara RB. 2005. Comparing the effects of genetic drift and fluctuating selection on genotype frequency changes in the scarlet tiger moth. Proc Roy Soc B 272: 211-217.

O'Quin CT, Drilea A, Roberts RB, Kocher TD. 2012. A small number of genes underlie male pigmentation traits in Lake Malawi cichlid fishes. J Exp Zool 318: 199-208.

Ohta T, Kimura M. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res 22: 201-204.

Olsson J, Eklöv P. 2005. Habitat structure, prey type and morphological reversibility: factors influencing phenotypic plasticity in perch. Evol Ecol Res 7: 1109-1123.

Ovenden J, Peel D, Street R, Courtney A, Hoyle S, et al. 2007. The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*), Mol Ecol 16: 127-138.

Owen RB, Crossley R, Johnson TC, Tweddle D, Kornfield I, Davison S, et al. 1990. Major low levels of Lake Malawi and their implications for speciation rates in cichlid fishes. Proc R Soc Lond B 240: 519-553.

Palm S, Laikre L, Jorde PE, Ryman N. 2003. Effective population size and temporal genetic change in stream resident brown trout (*Salmo trutta*, L.). Conserv Genet 4: 249-264.

Palstra FP, Fraser DJ. 2012. Effective/census population size ratio estimation: a compendium and appraisal. Ecol Evol 2: 2357-2365.

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

Parnell NF, Streelman JT. 2011. The macroecology of rapid evolutionary radiation. Proc Roy Soc B 278: 2486-2494.

Parsons KJ, Son YH, Albertson RC. 2011. Hybridization promotes evolvability in African cichlids: connections between transgressive segregation and phenotypic integration. Evol Biol 38: 306-315.

Patarnello T, Volckaert FAMJ, Castilho R. 2007. Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? Mol Ecol 16: 4426-4444.

Pauers MJ. 2011. One fish, two fish, red fish, blue fish: geography, ecology, sympatry, and male coloration in the Lake Malawi cichlid genus *Labeotropheus* (Perciformes: Cichlidae). Int J Evol Biol 2011: Article ID 575469.

Perry WL, Lodge DM, Feder JL. 2002. Importance of hybridization between indigenous and nonindigenous freshwater species: an overlooked threat to North American biodiversity. Syst Biol 51: 255-275.

Pfennig DW, Wund MA, Snell-Rood EC, Cruickshank T, Schlichting CD, Moczek AP. 2010. Phenotypic plasticity's impacts on diversification and speciation. Trends Ecol Evol 25: 459-467.

Piry S, Luikart G, Cornuet JM. 1999. Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90: 502–503.

Plenderleith M, van Oosterhout C, Robinson RL, Turner GF. 2005. Female preference for conspecific males based on olfactory cues in a Lake Malawi cichlid fish. Biol Lett 1: 411-414.

Podos J. 2001. Correlated evolution of morphology and vocal signal structure in Darwin's finches. Nature 409: 185-188.

Pollak E. 1983. A new method for estimating the effective population size from allele frequency changes. Genetics 104: 531-548.

Posada D. 2008. jModeltest: Phylogenetic model averaging. Mol Biol Evol 25:1253–1256.

Postl L, Herler J, Bauer C, Maderbacher M, Makasa L, Sturmbauer C. 2008. Geometric morphometrics applied to viscerocranial bones in three populations of the Lake Tanganyika fish *Tropheus moorii*. J Zool Syst Evol Res 46: 240-248.

Pray LA, Goodnight CJ, Stevens L, Schwartz JM, Yan G. 1996. The effect of population size on effective population size: an empirical study in the red flour beetle *Tribolium castaneum*. Genet Res 68: 151-155.

Price T. 1998. Sexual selection and natural selection in bird speciation. Phil Trans R Soc Lond B 353: 251-260.

Pritchard JK, Stephens M, Donelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Pritchard JR, Schluter D. 2001. Declining interspecific competition during character displacement. Evol Ecol Res 3: 209-220.

Raeymaekers JAM, van Houdt JKJ, Larmuseau MHD, Geldof S, Colckaert FAM. 2007. Divergent selection as revealed by P_{ST} and QTL-based F_{ST} in three-spined stickleback (*Gasterosteus aculeatus*) populations along a coastal-inland gradient. Mol Ecol 16: 891-905.

Rambaut A, Drummond AJ. 2009. Tracer v1.5. Available at http://beast.bio.ed.ac.uk/Tracer

Ray J, Husemann M, Danley PD, King RS. 2012. Genetic analysis reveals dispersal of Florida Bass haplotypes from reservoirs to rivers in Central Texas. Trans Am Fish Soc 141: 1269–1273.

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86: 248-249.

Reed DH, Frankham R. 2003. Correlation between fitness and genetic diversity. Conserv Biol 17: 230–237.

Reed DH, Briscoe DA, Frankham R. 2002. Inbreeding and extinction: The effect of environmental stress and lineage. Conserv Genet 3: 301–307.

Regan CT. 1922. The cichlid fishes of Lake Nyassa. Proc Zool Soc Lond 1921: 675-727.

Reinthal PN. 1990. The feeding habits of a group of herbivorous rock-dwelling cichlid fishes (Cichlidae: Perciformes) from Lake Malawi, Africa. Environ Biol Fish 27: 215-233

Ribbink AJ, Marsh BA, Marsh AC, Ribbink AC, Sharp BJ. 1983. A preliminary survey of the cichlid fishes of rocky habitats in Lake Malawi. S Afr J Zool 18: 149-309.

Ricketts TH. 2001. The matrix matters: effective isolation in fragmented landscapes. Am Nat 158: 87-99.

Rico C, Turner GF. 2002. Extreme microallopatric divergence in a cichlid species from Lake Malawi. Mol Ecol 11: 1585-1590

Rieseberg LH, Archer MA, Wayne RK. 1999. Transgressive segregation, adaptation and speciation. Heredity 83: 363-372.

Rieseberg LH, Widmer A, Arntz AM, Burke B. 2003a. The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. Phil Trans Roy Soc B 358: 1141-1147.

Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. 2003b. Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301: 1211-1216.

Roberts RB, Ser JR, Kocher TD. 2009. Sexual conflict resolved by the invasion of a novel sex determiner in Lake Malawi cichlid fishes. Science 326: 998-1001.

Rodriguez RL. 2012. Grain of environment explains variation in the strength of genotype x environment interaction. J Evol Biol 25: 1897-1901.

Roff DA, Fairbain DJ. 2012. A test of the hypothesis that correlational selection generates genetic correlations. Evolution 66: 2953-2960.

Rohlf JF. 2005. tpsRegr. Available at: http://life.bio.sunysb.edu/morph/.

Rohlf JF. 2006. tpsDig. Available at: http://life.bio.sunysb.edu/morph/.

Rohlf JF. 2007. tpsRelw. Available at: http://life.bio.sunysb.edu/ morph/.

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. Mol Ecol Notes 4: 137-138.

Rundell RJ, Price TD. 2009. Adaptive radiation, nonadaptive radiation, ecological speciation, nonecological speciation. Trends Ecol Evol 24: 394-399.

Sæther SA, Fiske P, Kålås JA, Kuresoo A, Luigujõe L, Piertney SB, Sahlman T, Höglund J. 2007. Inferring local adaptation from $Q_{ST} – F_{ST}$ comparisons: neutral genetic and quantitative trait variation in European populations of great snipe. J Evol Biol 20: 1563-1576.

Salzburger W. 2009. The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. Mol Ecol 18: 169-185

Salzburger W, Meyer A. 2004. The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. Naturwissenschaften 91: 277-290.

Salzburger W, S Baric S, Sturmbauer C. 2002. Introgressive hybridization in East African cichlids? Mol Ecol 11: 619-625.

Salzburger W, Mack T, Verheyen E, Meyer A. 2005. Out of Tanganyika: Genesis, explosive radiation, key-innovations and phylogeography of the Haplochromine cichlid fishes. BMC Evol Biol 5:17.

Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, et al. 2007. Ecological and evolutionary insights from species invasions. Trends Ecol Evol 22: 465-771.

Schander C, Halanych KM. 2003. DNA, PCR and formalinized animal tissue – a short review and protocols. Org Div Evol 3: 195–205.

Scheiner SM, Lyman RF. 1989. The genetics of phenotypic plasticity, I. Heritability. J Evol Biol 2: 95-107.

Schluter D, Clifford EA, Nemethy M, McKinnon JS. 2004. Parallel evolution and inheritance of quantitative traits. Am Nat 163: 809-822.

Schluter D, Grant PR. 1984. Determinants of morphological patterns in communities of Darwin's finches. Am Nat 123: 175-196.

Schluter D, McPhail JD. 1992. Ecological character displacement and speciation in sticklebacks. Am Nat 140: 85-108.

Schluter D. 2000. The ecology of adaptive radiation. Oxford Series in Ecology and Evolution, Oxford University Press, Oxford.

Schluter D. 2001. Ecology and the origin of species. Trends Ecol Evol 16: 372-380.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotech 18: 233-234.

Schwartz MK, Tallmon DA, Luikart G. 1998. Review of DNA-based census and effective population size estimators. Anim Conserv 1: 293-299.

Schwarzer J, Swartz ER, Vreven E, Snoeks J, Cotterill FPD, Misof B, Schliewen UK. 2012. Repeated trans-watershed hybridization among Haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution. Proc Roy Soc B 279: 4389-4398.

Scott LE, Johnson JB. 2010. Does sympatry predict life history and morphological diversification in the Mexican livebearing fish *Poeciliopsis baenschi*? Biol J Linn Soc 100: 608-618.

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW. 2011. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Mol Ecol Res 11: 1-8.

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

Seehausen O. 2013. Conditions when hybridization might predispose populations for adaptive radiation. J Evol Biol 26: 279-281.

Seehausen O, Van Alphen JJM. 1998. The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). Behav Ecol Sociobiol 42: 1-8.

Seehausen O. 2004. Hybridization and adaptive radiation. Trends Ecol Evol 19: 198-207.

Seehausen O. 2006. African cichlid fish: a model system in adaptive radiation research. Proc Roy Soc B 273: 1987-1998.

Seehausen O, Mayhew PJ, van Alphen JJM. 1999. Evolution of color patterns in East African cichlid fish. J Evol Biol 12: 514-534.

Selkoe T, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9: 615-629.

Servedio MR, van Doorn GS, Kopp M, Frame AM, Nosil P. 2011. Magic traits in speciation: 'magic' but not rare? Trends Ecol Ecol 26: 389-397.

Sharma R, Stuckas H, Bhaskar R, Khan I, Goyal SP, Tiedemann R. 2011. Genetically distinct population of Bengal tiger (*Panthera tigris tigris*) in Terai Arc Landscape (TAL) of India. Mamm Biol 76: 484-490.

Shaw KL. 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc Nat Acad Sci 99: 16122-16127.

Sherbakov DY. 1999. Molecular phylogenetic studies on the origin of biodiversity in Lake Baikal. Trends Ecol Evol 14: 92-95.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. Science 236: 787-792.

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

Smith PF, Kornfield I. 2002. Phylogeography of Lake Malawi cichlids of the genus *Pseudotropheus*: significance of allopatric colour variation. Proc Roy Soc Lond B 269: 2495-2502.

Spitze K. 1993. Population structure in *Daphnia obtusa* – quantitative genetic and allozymic variation. Genetics 135: 367-374.

Spreitzer ML, Mautner S, Makasa L, Sturmbauer C. 2012. Genetic and morphological population differentiation in the rock-dwelling and specialized shrimp-feeding cichlid fish species *Altolamprologus compressiceps* from Lake Tanganyika, East Africa. Hydrobiologia 682: 143-154.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate J. 2010. Adaptation genomics: the next generation. Trends Ecol Evol 25: 705–712.

Stauffer JR, van Snik Gray E. 2004. Phenotypic plasticity: its role in trophic radiation and explosive speciation in cichlids (Teleostei: Cichlidae). Anim Biol 54: 137-158.

Stauffer JR, Black K, Konings AF. 2013. Descriptions of five new species of *Metriaclima* (Teleostei: Cichlidae) from Lake Malawi. Zootaxa 3647: 101-136.

Stauffer JR, Bowers NJ, Kellogg KA, McKaye KR. 1997. A revision of the Blue-Black *Pseudotropheus zebra* (Teleostei: Cichlidae) complex from Lake Malawi, Africa, with description of a new genus and ten new species. Proc Acad Nat Sci Phil 148: 189-230.

Stearns SC. 1989. Evolutionary significance of phenotypic plasticity. BioScience 39: 436-445.

Stelkens RB, Seehausen O. 2009. Genetic distance between species predicts novel trait expression in their hybrids. Evolution 63: 884-897.

Stelkens RB, Schmid C, Selz O, Seehausen O. 2009. Phenotypic novelty in experimental hybrids is predicted by the genetic distance between species of cichlid fish. BMC Evol Biol 9: 283.

Strange JP, Knoblett J, Griswold T. 2009. DNA amplification from pin-mounted bumble bees (*Bombus*) in a museum collection: effects of fragment size and specimen age on successful PCR. Apidology 40, 134-139.

Streelman JT, Gmyrek SL, Kidd MR, Kidd C, Robinson RL, Hert E, Ambali AJ, Kocher TD. 2004. Hybridization and contemporary evolution in an introduced cichlid fish from Lake Malawi National Park. Mol Ecol 13: 2471-2479.

Streelman JT, Danley PD. 2003. The stages of vertebrate evolutionary radiation. Trends Ecol Evol 18: 126-131.

Streelman JT, Albertson RC, Kocher TD. 2007. Variation in body size and trophic morphology within and among genetically differentiated populations of the cichlid fish, *Metriaclima zebra*, from Lake Malawi. Freshwater Biol 52: 525-538.

Stuart YE, Losos JB. 2013. Ecological character displacement: glass half full or half empty? Trends Ecol Evol 28: 402-408.

Sturmbauer C. 2008. The Great Lakes in East Africa: biological conservation considerations for species flocks. Hydrobiologia 615: 95-101.

Sturmbauer C. 1998. Explosive speciation in cichlid fishes of the African Great Lakes: a dynamic model of adaptive radiation. J Fish Biol 53: 18-36.

Sturmbauer C, Baric S, Salzburger W, Rüber L, Verheyen E. 2001. Lake level fluctuations synchronize genetic divergence of cichlid fishes on African Lakes. Mol Biol Evol 18: 144-154.

Sturmbauer C, Husemann M, Danley P. 2011. New insights on explosive speciation and adaptive radiation from African cichlid fishes. In: Habel JC, Zachos FE: Hotspots, Springer.

Tagliavia M, Massa B, Albanese I, La Farina M. 2011. DNA extraction from Orthoptera museum specimens. Analyt Lett 44: 1058-1062.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 28: 2731-2739.

Thomas L, Buckland ST, Rexstad EA, Laake JL, Strindberg S, Hedley SL, Bishop JRB, Marques TA, Burnham KP. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. J Appl Ecol 47: 5-14.

Thomas L, Laake JL, Rexstad E, Strindberg S, Marques FFC, Buckland ST, Borchers DL, Anderson DR, Burnham KP, Burt ML, Hedley SL, Pollard JH, Bishop JRB, Marques TA. 2009. Distance 6.0. Release "2". Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance.

Tobler M, Carson EW. 2010. Environmental variation, hybridization, and phenotypic diversification in Cuatro Ciénegas pupfishes. J Evol Biol 23: 1475-1489.

Tobler M, DeWitt TJ, Schlupp I, García de León FJ, Herrmann R, Feulner PGD, Tiedemann R, Plath M. 2008. Toxic hydrogen sulfide and dark caves: phenotypic and genetic divergence across two abiotic environmental gradients in *Poecilia mexicana*. Evolution 62: 2643-2659.

Tobler M, Palacios M, Chapman LJ, Mitrofanov I, Bierbach D, Plath M, et al. 2011. Evolution in extreme environments: replicated phenotypic differentiation in livebearing fish inhabiting sulfidic springs. Evolution 65: 2213-2228.

Torres-Dowdall J, Handelsman CA, Reznick DN, Ghalambor CK. 2012. Local adaptation and the evolution of phenotypic plasticity in Trinidadian guppies (*Poecilia reticulata*). Evolution 66:3432-3443.

Traill LW, Bradshaw CJA, Brook BW. 2007. Minimum viable population size: a meta-analysis of 30 years of published estimates. Biol Conserv 139: 159-166.

Turner TF, Wares JP, Gold JR. 2002. Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). Genetics 162: 1329-1339.

Valqui J, Hartl GB, Zachos FE. 2010. Non-invasive genetic analysis reveals high levels of mtDNA variability in the endangered South-American marine otter (*Lontra felina*). Conserv Genet 11: 2067-2072.

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

Van Oppen MJH, Turner GF, Rico C, Robinson RL, Deutsch JC, Genner MJ, Hewitt GM. 1998. Assortative mating among rock-dwelling cichlid fishes supports high estimates of species richness from Lake Malawi. Mol Ecol 7: 991-1001.

van Staaden MJ, Smith AR. 2011. Cutting the Gordian knot: Complex signaling in African cichlids is more than multimodal. Curr Zool 57: 237-252.

Vandermeer J, Evans MA, Foster P, Höök T, Reiskind M, Wund M. 2002. Increased competition may promote species coexistence. Proc Nat Acad Sci 99: 8731-8736.

Via S, Lande R. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39: 505-522.

Via S, Gomulkiewicz R, de Jong G, Scheiner SM, Schlichting CD, van Tienderen PH. 1995. Adaptive phenotypic plasticity: consensus and controversy. Trends Ecol Evol 10: 212-217.

Vokey JE, Burton D. 1998. Responsiveness to noradrenaline of melanophores associated with cryptic patterning in winter flounder, *Pleuronectes americanus*, in vitro. Can J Zool 76: 1837-1841.

Vucetich JA, Waite TA, Nunney L. 1997. Fluctuating Population Size and the Ratio of Effective to Census Population Size. Evolution 51: 2017-2021.

Waits ER, Bagley MJ, Blum MJ, McCormick FH, Lazorchak JM. 2008. Source-sink dynamics sustain central stonerollers (*Campostoma anomalum*) in a heavily urbanized catchment. Freshw Biol 53: 2061-2075.

Walker JA. 2010. An integrative model of evolutionary covariance: a symposium on body shape in fishes. Integr Comp Biol 50: 1051-1056.

Wandeler P, Hoeck PEA, Keller LF. 2007. Back to the future: museum specimens in population genetics. Trends Ecol Evol 22: 634– 642.

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

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

Waples RS. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered 89: 438-450.

Waples RS. 2005. Genetic estimates of contemporary effective population size: to what time periods do estimates apply? Mol Ecol 14: 3335-3352.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358-1370.

West-Eberhard MJ. 1989. Phenotypic plasticity and the origins of diversity. Ann Rev Ecol Syst 20: 249-278.

Whitlock MC, Guillaume F. 2009. Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. Genetics 183: 1055-1063.

Whitlock MC. 2000. Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. Evolution 54: 1855-1861.

William EG, Slatkin M. 1999. Using maximum likelihood to estimate population size from temporal changes in allele frequencies. Genetics 152: 755-761.

Willis SC, Winemiller KO, López-Fernández H. 2005. Habitat structural complexity and morphological diversity of fish assemblages in a Neotropical floodplain river. Oecologia 142: 284-295.

Wimberger PH. 1992. Plasticity of fish body shape. The effects of diet, development, family and age in two species of *Geophagus* (Pisces: Cichlidae). Biol J Linn Soc 45: 197-218.

Won Y-J, Sivasundar A, Wang Y, Hey J. 2005. On the origin of Lake Malawi cichlid species: A population genetic analysis of divergence. Proc Nat Acad Sci 102: 6581-6586.

Wright S. 1951. The genetical structure of populations. Ann Eug 15: 323–354.

Zachos FE, Althoff C, Steynitz Y, Eckert I, Hartl GB. 2007. Genetic analysis of an isolated red deer (*Cervus elaphus*) population showing signs of inbreeding depression. Euro J Wildlife Res 53: 61-67.

Zelditch M, Swiderski D, Sheets H, Fink W. 2004. Geometric morphometrics for biologists. Elsevier Academic Press, Amsterdam.

Zidana H, Turner GF, van Oosterhout C, Hänfling B. 2009. Elevated mtDNA diversity in introduced populations of *Cynotilapia afra* (Günther 1894) in Lake Malawi National Park is evidence for multiple source populations and hybridization. Mol Ecol 18 4380-4389.

Zimmerman EC. 1970. Adaptive radiation in Hawaii with special reference to insects. Biotropica 2: 32-38.