EVOLUTION OF LAKE MALAWI CICHLID FISHES
(PERCIFORMES: TELEOSTEI)

By

Peter F. Smith

B.A. University of Maine, 1997

A THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
(in Biological Sciences)

The Graduate School
The University of Maine
December, 2002

Advisory Committee:

Irv Kornfield, Professor of Zoology, Advisor

William Glanz, Associate Professor of Zoology

Thomas Kocher, Professor of Zoology and Genetics, University of New Hampshire

Paul Rawson, Assistant Professor of Marine Science

Mary Tyler, Professor of Zoology

EVOLUTION OF LAKE MALAWI CICHLID FISHES

(PERCIFORMES: TELEOSTEI)

By Peter F. Smith

Thesis Advisor: Dr. Irv Kornfield

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Biological Sciences)
December, 2002

The cichlid fish species flocks of East Africa provide a compelling model system in which to study the process of speciation. In Lake Malawi, greater than 1000 species of cichlids have emerged since the filling of the lake basin about 1 million years ago. Over 99% of the Lake Malawi haplochromine cichlids are endemic, suggesting that most of this diversification has taken place within the temporal and spatial boundaries set by the Lake's shores. Moreover, many of these species are endemic to small areas within the lake, indicating that speciation has occurred very recently or perhaps is in progress in some instances. Examination of the genetic structure of recently divergent populations or species is relevant to our understanding of the process of speciation in general.

Here I present work on several populations, insipient species and species of Lake Malawi haplochromine cichlids. I have examined the genetic diversity of microsatellite DNA markers and applied these data to specific hypotheses relating to speciation. First I review the literature relevant to the latter chapters. Chapter two is an examination of several populations with two distinct color patterns and lake-wide distributions. This chapter contributes genetic data to a persistent biogeographical problem among these

fishes – monophyly of color pattern versus independent derivation of similar patterns in disparate portions of the lake. This work showed that many like-colored populations share genetic similarity, but one population likely represents parallel evolution of coloration. In chapter 3, I present genetic data that suggests that hybridization has occurred among closely related species differing in coloration. These data contribute to a growing list of confirmed and suspected cases of hybridization among African cichlids, and is the first genetic data supporting hybridization among congeners in Lake Malawi. Finally, I present a microsatellite phylogeny of several populations and newly derived species. This examination confirmed conclusions from chapter two, that color convergence is a common phenomenon among Lake Malawi cichlids. Indeed several characters of coloration evolved multiple times among the taxa examined. This dissertation contributes several new findings to the field of speciation among African cichlids, though many questions remain. While it is unlikely that a single mechanism is responsible for generating this remarkable species diversity, revealing the conditions that have allowed the cooperation of multiple mechanisms will be a fascinating advance in the study of evolution.

ACKNOWLEDGEMENTS

I thank the members of my dissertation committee, Drs. Irv Kornfield, Bill Glanz, Tom Kocher, Paul Rawson and Mary Tyler. Their willingness to sacrifice their time and energy in order to help me reach my goals is a tribute to their vocation as educators. I am especially thankful to Irv Kornfield, who offered me an exceptional environment of guidance and freedom that has facilitated my progress.

I am also grateful to Ad Konings who participated in the fieldwork portion of my project. He has consistently made himself readily available to answer all of my questions about the fascinating cichlids of Lake Malawi, and has been completely unselfish with his extensive knowledge and his many thousands of underwater photographs. Several times throughout this project Ad has offered important insights that have helped to advance my work. Thanks also to Martin Geerts for helping in the field and for contributing some important bibliographic references.

I am also grateful to the many cichlid researchers who have preceded me and those who are currently working on problems in cichlid evolutionary biology. All of the work in this dissertation depends heavily on the work of many other scientists.

Finally I extend my gratitude to the family and friends who have supported me throughout my years of education. Most notably, I thank my lovely wife, Ashley, who has been a dedicated partner. She has provided abundant love and encouragement; I am endlessly grateful.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS……ii

LIST OF TABLES……vii

LIST OF FIGURES……viii

Chapter

1. LITERATURE REVIEW AND INTRODUCTION……1

- Abstract……1
- Introduction……1
- Species concepts, diversity and taxonomy……4
  - Species definitions……5
  - Taxonomic diversity……6
- Centrality of phylogenetic reconstruction……7
- Geographical models of speciation……11
  - Allopatry……12
  - Evolutionary implications of lake level fluctuations……13
  - Sympatry……16
- Application of sympatric speciation scenarios to the East African Great Lakes……18
- Mechanisms driving differentiation……19
  - How important is local adaptation to species diversity?……21
  - Is genetic drift critical to diversification?……22

- What role might hybridization play in faunal evolution?....................23
- Is sexual selection driving cichlid speciation?..............................25
    - Proposed mechanisms of divergence by sexual selection...........25
    - Appeal of speciation by sexual selection.............................26
    - Testing divergence hypotheses.......................................27

Evolutionary generalizations and future studies....................................29
- Understanding the genetic basis of relevant traits............................30
- Understanding coloration in cichlids...30
- Studying behavior...31
- Insights from population genetics...31
- Criteria for evaluating models of divergence 32

2. PHYLOGEOGRAPHY OF LAKE MALAWI CICHLIDS OF THE GENUS METRIACLIMA: SIGNIFICANCE OF ALLOPATRIC COLOR VARIATION......33

- Abstract..33
- Introduction..34
- Materials and methods...40
    - Study species...40
    - Sampling & DNA preparation..42
    - Data analysis...42
- Results...44

Microsatellite loci.....44

Genetic relationships.....45

Discussion.....48

3. HYBRID ORIGIN OF A CICHLID POPULATION IN LAKE MALAWI: IMPLICATIONS FOR SPECIES DIVERSITY AND GENETIC VARIATION.....54

Abstract.....54

Introduction.....54

Materials and methods.....60

Results.....61

Discussion.....70

4. EXTREME COLOR CONVERGENCE WITHIN THE *Pseudotropheus (Maylandia) zebra* SPECIES COMPLEX OF LAKE MALAWI CICHLIDS IS REVEALED BY A MICROSATELLITE PHYLOGENY.....76

Abstract.....76

Introduction.....77

Materials and methods.....79

Study species.....79

Sampling & DNA preparation.....80

Data analysis.....83

Results.....84

Discussion.....87

Conclusions.....91

LITERATURE CITED..92

BIOGRAPHY OF THE AUTHOR..114

LIST OF TABLES

Table 1.1. Models of cichlid speciation..20

Table 2.1 Genetic diversity of M. zebra populations examined..............................41

Table 2.2 Pairwise Fst (and standard deviations) for *M. zebra* population comparisons...46

Table 3.1 Genetic diversity of populations examined...62

Table 3.2 Allele frequencies for three populations and the simulated hybrid sample at four microsatellite loci...64

Table 3.3 Exact tests of population differentiation by locus...................................66

Table 4.1 Location and genetic diversity of population samples............................81

LIST OF FIGURES

Figure 1.1. Intrapopulation color diversity within the Lake Malawi *Metriaclima zebra* species complex. 8

Figure 1.2. The effects of repeated lake fluctuations in cichlid species diversity.........15

Figure 2.1 Examples of RT an BB phenotypes and graphical representation of the hypotheses tested...35

Figure 2.2. Diagram of Lake Malawi showing seven *M. zebra* and two *M. sp.* 'zebra Chilumba' sampling locations. 37

Figure 2.3. Neighbor-joining, majority rule consensus tree of the seven sampled populations using Nei's standard distance (Ds) calculated from nine microsatellite DNA markers. 47

Figure 2.4. Factorial Correspondence Analysis (FCA) shows strong separation of M. sp. 'zebra Chilumba' individuals from the southern BB (Namalenje Island) and RT (Namaso Bay) individuals.........................49

Figure 3.1. A hybrid population in southeastern Lake Malawi. 59

Figure 3.2. Genotype assignment test (four loci) of all samples to putative parental baselines (*M. thapsinogen* and *M. zebra*). 67

Figure 3.3. Graphical representation of genetic variance among samples by factorial correspondence analysis (FCA). 68

Figure 4.1. Sampling locations for seventeen populations with five distinct color phenotypes from ten species of the *Pseudotropheus (Maylandia) zebra* species complex. 82

Figure 4.2. An unrooted phylogeny for the seventeen taxa in the Pseudotropheus (Maylandia) zebra species complex showing extensive convergence in color phenotype……………………………………85

Chapter 1

LITERATURE REVIEW AND INTRODUCTION

Abstract

Cichlid fishes (Perciformes: Teleostei) found in the lakes of Africa have served as model systems for the study of evolution. The enormous number of species present (1000 in Lake Malawi alone), the great diversity of trophic adaptations and behaviors, and the extreme rapidity of their divergence (< 50,000 yrs for some faunas) single out these organisms as examples of evolution in progress. Because these fishes are confined to discrete lacustrine environments and their emergence is bounded by geological features, these groups provide platforms in which to study evolution. We review theoretical studies and empirical research on the cichlid faunas of Africa to provide a synthetic overview of current knowledge of the evolutionary processes at work in the Cichlidae. This view provides the critical information needed to formulate and test hypotheses which may permit discrimination among the diverse theories and models that have been advanced to explain the evolution of these fishes.

Introduction

Periods of explosive speciation and adaptive radiation have punctuated the history of metazoan evolution. The cichlid fish faunas of lakes in Africa present singular opportunities to study such events. Lineages within the Cichlidae have diversified extensively in both the New World and the Old World (Keenleyside 1991), but it is in Africa where the most spectacular examples of radiation and speciation are found.

Because of this, African cichlids have been incorporated into many basic evolutionary texts (Strickberger 2000).

The cichlid fishes of lakes Malawi, Tanganyika, Victoria (Goldschmidt 1996, Kawanabe et al. 1997, Martens et al. 1994), and the other, much smaller African lakes (Fryer and Iles 1972, Schliwewen et al. 1994, Seegers and Tichy 1999) provide unique opportunities with which to investigate evolutionary processes (Echelle and Kornfield 1984, Keenlyside 1991, Ricklefs and Schluter 1993). Time scales associated with diversification for many of these complexes are extremely short (Kornfield 1978) (< 50 kyr for some faunas), the extent of trophic differentiation is extensive (Fryer and Iles 1972) (ranging from plankton grazers to egg predators), and the number of endemic taxa is enormous (1000 species in Lake Malawi is a reasonable estimate (Turner 1994)). Because of these factors, the cichlids of African lakes provide geologically bounded systems within which to explore fundamental processes of speciation and diversification.

There is an extremely rich, speculative literature concerning the mechanisms underlying the processes of cichlid divergence (Galis and Metz 1998, Meyer 1993, Staiassny and Meyer 1999, Turner 1997). The rapidity of radiations has suggested to many that some exceptional mechanism(s), including special water characteristics (Ladiges 1968), habitat complexity (Temple 1969), lake age (Ladiges 1968), lake level fluctuation (Fryer 1959, Trewavas 1935), predation (Greenwood 1965), mutation (Fryer and Iles 1972), hybridization (Crapon de Caprona and Fritzsch 1984), reproductive characteristics (Dominy 1984, Kosswig 1947), and trophic polymorphism (Sage and Selander 1975) are necessary to explain the extent of cichlid radiation. Similarly, allopatric (Greenwood 1965, McKaye and Gray 1984), microallopatric (Fryer 1959), and sympatric (Kosswig 1947, Schliwewen et al. 1994, Seehausen and Van Alphen 1999, Turner 1994) modes of speciation have been suggested. There is general consensus that almost all of the endemic cichlid species and genera of Malawi and Victoria originated within the lakes proper (Mayr 1984); intralacustrine speciation is extensive in Tanganyika, though this lake was clearly seeded several times (Sturmbauer et al. 1997). However, despite a wealth of recent theoretical and modeling efforts, limited empirical progress has been made in determining the critical factors that have accelerated or retarded diversification.

The possibility of sexual selection has raised enormous interest (Turner 1997), but there have been few critical examinations of more conventional models of diversification. That is, alternate population processes such as subdivision, isolation and drift may be significant driving forces for speciation at the geographic level. The extent of panmixia is not known, and the possibility of geographic structuring on local or regional scales has received limited attention (see below). Similarly, tests of many hypotheses have been hampered by the absence of reliable phylogenies for closely related species. It may be beneficial to seek to reject null hypotheses of spatial variation prior to proposing more elaborate explanations for diversification.

Here we assemble and review basic data on the geological and biological attributes that make African cichlid species unique and useful systems in which to study evolution. These faunas challenge simple application of species concepts; clarification of and consensus on this topic is of central importance to resolve persistent questions in cichlid evolution. Geographical models of speciation in African cichlids are summarized and empirical studies that have been offered as support for specific models are critiqued.

The effects of fluctuating lake levels as a promoter of allopatric divergence are presented and examined. We offer a critical evaluation of the mechanisms potentially contributing to this spectacular diversification, discuss classical evolutionary forces as they apply to these cichlid fish systems, and offer speculation on less established factors such as hybridization and the emerging mechanism of favor: sexual selection.

Species concepts, diversity and taxonomy

Discussions about definitions of species have dominated the cichlid literature for many years. Though most work before 1980 assumed the Biological Species Concept (BSC) of Mayr, subsequent studies considered alternative definitions. While the BSC or "isolation concept" has been called the most useful species concept for the study of speciation (Coyne and Orr 1998), alternative approaches have been explicitly championed by a number of cichlid researchers. The recognition concept of Patterson (1985), which defines species as groups that share common fertilization systems – or specific mate recognition systems, and the cohesion concept (Templeton 1989) have been emphasized (Greenwood 1991, Ribbink 1986, Ribbink 1994, Sturmbauer 1998, Turner 1994).

Understanding the speciation process may depend upon the species concept employed (Templeton 1989). For this reason, several authors have chosen non-relational species concepts (Ribbink 1994, Sturmbauer 1998, Turner 1994) in the study of cichlid evolution. However, this debate can also be viewed as semantic. Indeed, regardless of the species concept used, the processes (genetic drift, natural selection, sexual selection, etc.) involved in the early divergence of populations are the same, whether they are seen as divisive or cohesive. Confusing issues surrounding the use of the BSC may result from what has been observed as a tendency of workers to create caricatures of one species concept in order to clarify the reasoning of another (Harrison 1998). Here we adhere to the reasoning of the BSC (the isolation concept), but do not reject the incorporation of alternate views.

Species definitions

How are species recognized? Because of their extreme youth, patchy distribution, and phenotypic similarity, distinguishing species is problematic (Sturmbauer 1998). When sympatric, biological species have been demonstrated by in situ observations of assortative mating by color morphs (eg, Holzberg 1978, Seehausen et al. 1998).More typically, species have been established (or their status confirmed), by studies that report significant differences in frequencies of molecular markers among sympatric morphs (Kornfield 1978, McKaye et al. 1982, McKaye et al. 1984, VanOppen et al. 1998) or by laboratory examination of mate discrimination (Knight and Turner 1999, Knight et al. 1998, Seehausen et al. 1997).

In allopatry, differences in male breeding coloration have been uniformly employed to recognize species, while qualitative differences in dentition and osteology have been important in defining genera and species groups (Lewis 1982). It is nearly universally presumed that the existence of unique male breeding coloration is sufficient to delimit species (Eccles and Trewavas 1989, Greenwood 1994, Konings 2001, Lewis 1982, Ribbink et al. 1983, Turner 1994). Though a few taxa have wide distributions, most species defined on the basis of color are 'narrow endemics', restricted to specific islands or rocky outcrops (Konings 2001, (Ribbink et al. 1983). Note that under an ecological species concept, closely-related allopatric populations differing only in color may be the same species if they all occupy the same ecological niche. Indeed, many geographic variants, strikingly different in coloration, possess indistinguishable trophic adaptations and foraging ecologies (Sturmbauer and Meyer 1992). Such geographical color variants may be of central importance to reproductive isolation and generation of diversity. Comparative studies of sympatric and allopatric populations in different stages of divergence may provide the most profound insights into geographic modes of speciation (Foster et al. 1998).

Taxonomic diversity

Most species of African cichlids have yet to be formally described. The task of cataloguing the diversity of the East African great lakes is particularly difficult both because of the large number of species (Goldschmidt 1996) and by the obscurity of species boundaries (Sturmbauer 1998). In the case of allopatric populations showing geographic color variation, reproductive isolation may never be tested in nature. While they are not necessarily biological species, the magnitude of novel coloration may merit taxonomic recognition. For example, newly discovered cichlids are given descriptive names in order to distinguish them from all other populations (i.e., "Pseudotropheus species tropheops mbenji blue" see Konings 2001). With extremely few exceptions, such allopatric variants have not been assigned subspecific epithets, although several such populations have recently been given specific status (Stauffer et al. 1997). However, because of the ubiquity of geographic variation, considerable care needs to be exercised in defining taxa and interpreting taxonomic definitions. In particular, we emphasize the need to examine adequate sample sizes of organisms and to make in situ observations of coloration; in the absence of such practices, artificial species may be created (Fig. 1.1, Stauffer et al. 1997). Additionally, it is important to recognize the subtle yet critical difference between population differentiation and speciation (Magurran 1998)). The inability to distinguish between these two biological processes, and between types of species (eg, biological vs. phylogenetic, (Harrison 1998)), may confound evolutionary studies.

Centrality of phylogenetic reconstruction

Differentiating alternative explanations for patterns of divergence and the processes that drive them requires robust hypotheses of relationships among species. Organismal phylogenies are essential for adjudication of the roles that special processes or characters may play in the speciation process. In the Cichlidae, there is extensive potential for convergence or parallelism in trophic or morphological domains (Lazarro 1991). Among haplochromines, the most speciose lineage of African cichlids, morphological similarity and ecological equivalence are patent among the three major lakes (Fryer and Iles 1972, Stiassny 1981). Within lakes, the magnitude of similarity is striking as well (Eccles and Trewavas 1989, Greenwood 1974). If convergence is real, the use of standard ichthyological characters for the construction of phylogenies, particularly those associated with dentition, can be circular.

The study of lepidological variables (those derived from scale morphology and squamation pattern), pioneered by Lippitch (1993), has proved exceedingly valuable in

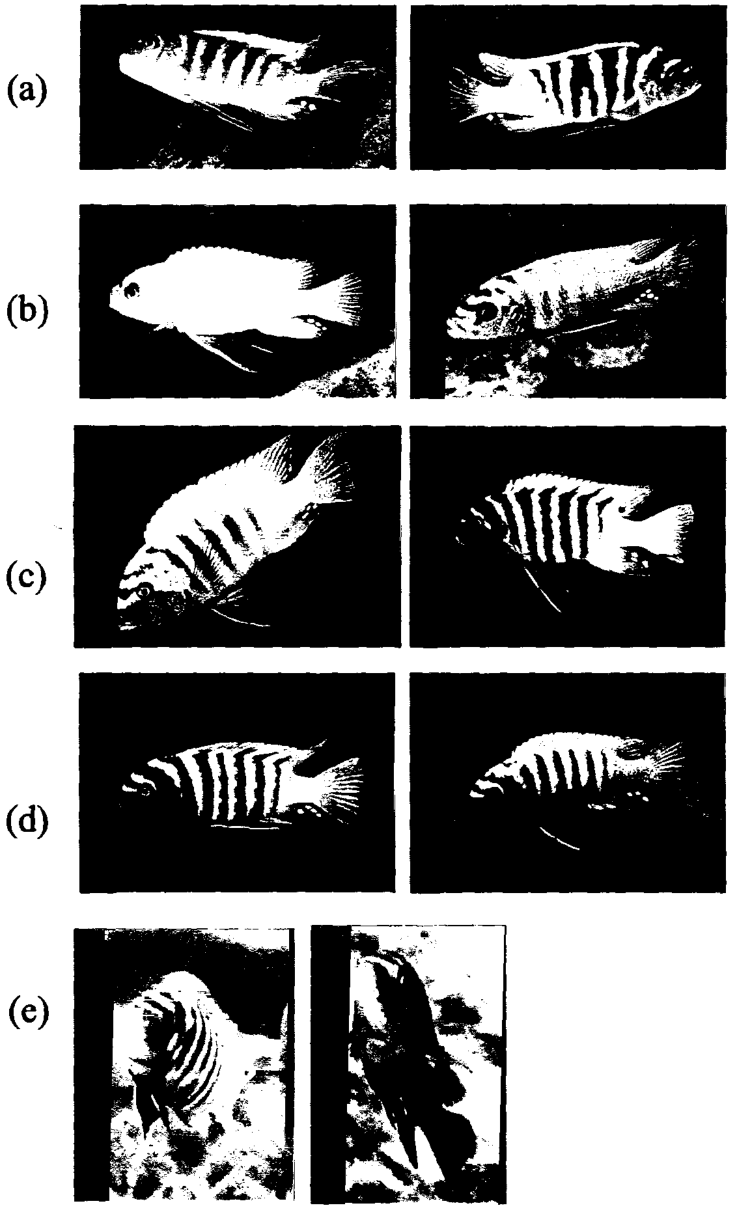

Figure 1.1. Intrapopulation color diversity within the Lake Malawi Metriaclima zebra species complex. Variation in the illustrated traits is continuous; presented are examples of extreme phenotypes. (a) Pattern assymetry (flanks of a single "red dorsal" individual - Chimwalani); (b) bar density (Mazinzi Reef - M.benetos); (c) dorsal fin bar penetration ("Blue Black" - Meponda Rocks); (d) dorsal fin bar penetration ("red dorsal" Nakantenga Island - M. pyrsonotos); (e) throat patch coloration (Eccles Reef - M. thapsinogen). (Photographs a-d by Ad. Konings).

establishing interrelationships among some groups of African cichlids. They have been used to demonstrate monophyly in the well-defined tribes in Lake Tanganyika and to erect a phylogeny among some of these lineages (Lippitch 1998); they also appear useful for defining lineages among some groups of more closely related haplochromines (Seehausen et al. 1998). While general application of this approach to some haplochromine taxa has been problematic (Kuusipalo 1998), it shows extreme promise for well-defined groups.

Given the potential for convergence, it is clear that phylogenies generated from characters independent of morphology are desirable. More specifically, findings from molecular studies are central for resolving debates involving such attributes as behavior and morphology. Molecular studies have clarified relationships for the Cichlidae (Farias et al. 1998), among African great lake cichlids, and have begun to provide insights among taxa within lakes. The seminal paper by Meyer and coworkers (Meyer et al. 1990) using mtDNA demonstrated monophyly of the faunas of Malawi and Victoria and established relationships of the cichlid faunas among lakes. The often cited convergence among the mbuna of Malawi and the rockfishes of Tanganyika was unambiguously confirmed by Kocher and colleagues (Kocher et al. 1993) with comparative mtDNA sequence analysis; subsequent analyses established sister-group relationships between Malawi and Tanganyika faunas (Meyer 1993). Relationships among tribes within Tanganyika and confirmation of multiple invasions of portions of this fauna were initially established with allozymes (Nishida 1991, Nishida 1997) and have been extended and clarified using mtDNA data (Kocher et al. 1995, Sturmbauer and Meyer 1993, Sturmbauer et al. 1994). SINES (Short Interspersed Nuclear Elements) have been exploited to support monophyly of several Tanganyikan tribes (Takahashi et al. 1998). The unique insertion events associated with SINES are compelling, but the sensitivity of this assay among very closely related cichlids may be limited.

Elucidating the relationships among more recently diverged taxa using molecular techniques has been problematic. While mtDNA proved useful to examine divergence within several lineages in Tanganyika (Sturmbauer and Meyer 1992), it discriminates only major, divergent clades in the Malawi fauna (Moran et al. 1994). The reason that mtDNA failed at lower levels is that many taxa retained ancestral polymorphisms (Moran and Kornfield 1993, Moran and Kornfield 1995); comprehensive analysis of mtDNA revealed extensive retained polymorphism throughout the Malawi fauna (Parker and Kornfield 1997). The absence of lineage sorting of mtDNA compromises interpretation of molecular data in a phylogenetic context (Bowers et al. 1994, Reinthal and Meyer 1997) where gene trees may masquerade as species trees (Avise 1994). In Victoria and surrounding waters, neutral nuclear polymorphisms have persisted in cichlids for extended periods of time (Nagl et al. 1998). Extreme caution needs to be exercised in exploiting sequence data (Klein et al 1993, Ono et al 1993, Sultmann and Mayer 1997) for phylogenetic analyses in cichlid fishes, particularly where samples sizes are limited. Note, however, that molecular polymorphisms may be useful tools for examining other aspects of the evolutionary process (Clark 1997, Klein et al. 1993) such as effective population size.

Multilocus comparisons have demonstrated the utility of microsatellites to distinguish groups of closely related taxa (Kornfield and Parker 1997). This class of markers appears to be sufficiently informative to erect statistically compelling phylogenies among some closely related species (PF Smith & I Kornfield, in preparation). In like manner, large numbers of AFLPs (Amplified Fragment Length Polymorphisms) have been used to examine phylogenetic relationships among several well defined groups of mbuna (Albertson et al. 1999). In distance analyses, significant phylogenetic affinities among groups of congeners were demonstrated, but statistical support for intergeneric relationships was more limited (Albertson et al. 1999).

In general, the difficulty in establishing phylogenies for such taxa reflects the short intervals of absolute time that must have transpired during generic-level divergence. If diversification were concentrated in short temporal periods associated with major hydrological events such as lake expansions, speciation may effectively occur over wide geographic areas. In such cases of near simultaneity, resultant star phylogenies may reflect hard polytomies (Hoelzer and Melnick 1994) rather than an inability to resolve underlying signals. Distinguishing such bursts of speciation from soft polytomies may be possible if there is resolution above and below the event in question (Lessa and Cook 1998). Because of the desirability of a phylogenetic framework for evaluating hypotheses of origination and their associated mechanisms (Cannatella et al. 1998), additional research should be vigorously pursued.

Geographical models of speciation

Monophyly of Lake Malawi and of Lake Victoria cichlids suggest intralacustrine speciation, but strictly speaking, this is different from sympatric speciation. Several authors have developed scenarios whereby intralacustrine speciation occurs in an allopatric mode. Macro-allopatric speciation results from differentiation between separate basins, and micro-allopatric speciation is caused by differentiation between patches of fragmented habitat or via isolation by distance (Meyer et al. 1994, Rossiter 1995). Sympatric speciation is used in reference to divergence in the absence of geographical barriers to gene flow (Seehausen and VanAlphen 1998). Here we consider speciation from allopatric and sympatric perspectives.

Allopatry

One general model for allopatric speciation, the "Nabugabo scenario", was articulated by Greenwood (1965) for Lake Victoria. In this widely cited scheme, very minor fluctuations in lake level would serve to isolate fish in small, peripheral lakes over short periods of time (<4000 yr); such lakes could generate endemics which might retain specific identity upon reunion with Victoria and thus contribute to cumulative diversity. Difficulties with the evaluation of this model based on the present condition of the lake (Kaufman and Ochumba 1993) have included the inability to confirm that extant "peripheral endemics" are truly absent from Lake Victoria and the observation that some "peripheral endemics" have been found in multiple satellite lakes (Fournieret et al. 1992, Ogutu-Ohwayo 1993). There has also been faunal interchange among other lakes in the region (Kaufman 1997). Nevertheless, given that the extensive and exceedingly complex shoreline of Lake Victoria has not been systematically surveyed, the idea that novel assemblages of endemics might have persisted and reinvaded Victoria remains appealing.

Fryer (1959) presented a hypothesis of intralacustrine speciation that does not require fragmentation of a lake basin into separate bodies of water. In his scenario for

Lake Malawi, allopatry is achieved by highly philopatric cichlids occupying isolated areas of available habitat, allowing genetic divergence to take place over distance (see Kornfield 1978, Markert et al. 1999). In Lake Tanganyika, secondary contact between divergent populations of *Tropheus* inferred from biogeography illustrated microallopatric divergence (Fryer and Iles 1972). Subsequent molecular studies of a number of Tanganyikan rockfishes (Sturmbauer et al. 1997) have revealed patterns ranging from localized differentiation to continuity. The extent of population structure is important to all allopatric models of cichlid speciation, regardless of the proposed mechanism, because greater population subdivision implies a greater number of discrete entities subject to genetic drift and selection. While divergence in allopatry (micro-allopatry) may provide some insight into the early stages of intra-lacustrine speciation, there is a great deal still unknown about why divergence of sexual characteristics in allopatry would be so rapid and how newly divergent populations could survive presumed repeated bouts of secondary contact to become new species.

Evolutionary implications of lake level fluctuations

Changes in available habitat are associated with lake level fluctuations; these events could be central to understanding broader patterns of radiation and, perhaps, mechanisms controlling diversity and rates of speciation. Repetitive changes in lake level might continually create new opportunities for isolation in allopatry. Such potential effects were explicitly noted in early studies (Fryer 1959, Greenwood 1965) and have been considered by many others (McKaye and Gray 1984, Sturmbauer 1998). For example, two hypotheses have been presented to explain the presence of many species endemic to small islands in the southern shallow areas of Lake Malawi. First, it may be that these taxa predate the most recent rise in lake level; in this scenario, current island endemics arose through colonization of the newly flooded islands by species followed by their subsequent extinction in their former ranges. Alternately, when the southern portion of the lake was flooded, the newly formed islands were initially colonized at random by individuals of various mbuna species that are still extant elsewhere in the lake; these founding individuals subsequently diverged to produce new taxa. Repeated lake-level fluctuation events could be viewed as facilitating taxon-cycles (Kaufman 1997) and, depending upon biogeography and lake basin topography, could act as "species pumps" (Rossiter 1995) to increase endemic diversity.

An appreciation of the details of this process offer opportunities to understand the roles that natural selection and genetic interchanges may play in the overall process (Fig. 1.2). In the most general model, discrete basins and their surrounding habitats would act as refugia during lake regressions. Extinction would be elevated during these periods; taxa that are resident in such areas would not be tested by intense competition to the same extent as species that were forced to move into them. Previously allopatric taxa that were trophically distinct, or those that exploited some novel aspect of reproduction would be candidates for persistence. During such ecological crunches, significant trophic differentiation might occur in some lineages, though overall species diversity should decrease. Reproductively isolated ecological cognates would be particularly susceptible to extinction. Narrow endemics which were uncommon or rare (Ribbink et al. 1983) would be expected to have high rates of extinction relative to taxa with large population sizes. In aggregate, diversity in refugia would increase as new taxa were added, but

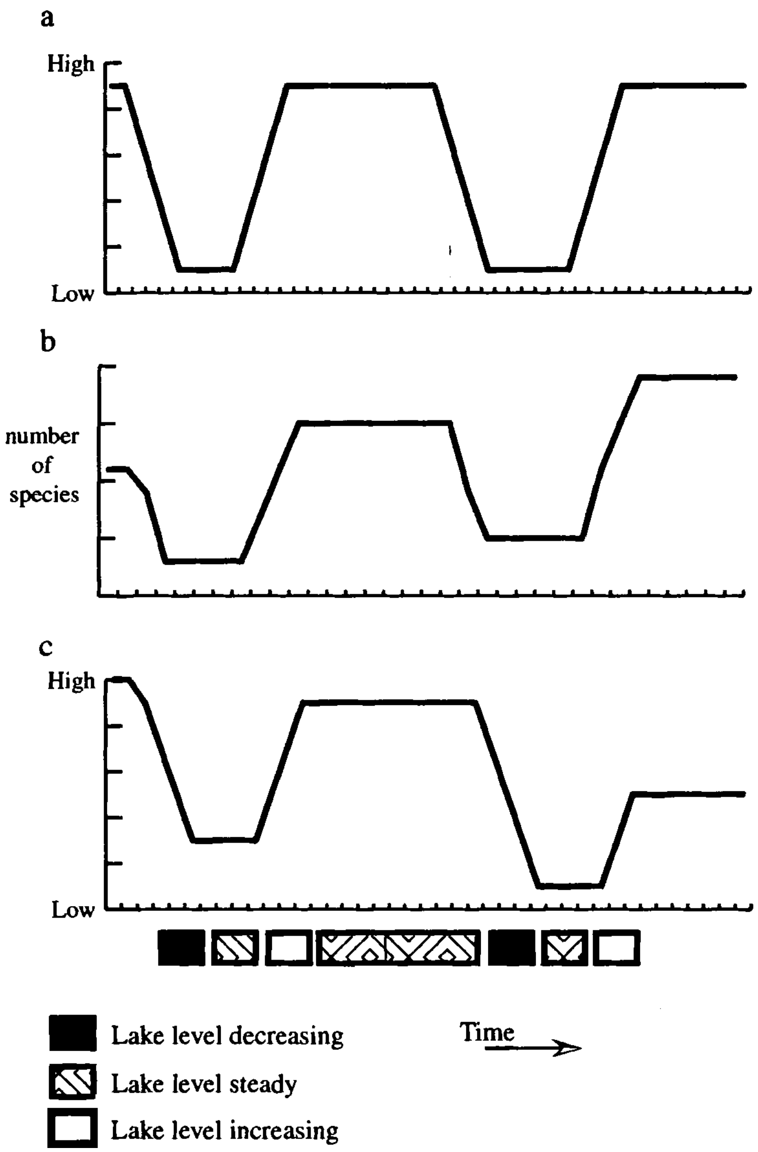

Figure 1.2. The effects of repeated lake fluctuations in cichlid species diversity. Extinction rates are high during periods of lake level drops (black), while speciation rates are high during lake level rises (white); when lake level is constant (hatched), speciation and extinction are in equilibrium. (a) Net number of species remains constant; (b) number of species increases with fluctuation cycles, the species pump (Rossiter 1995); (c) number of species decreases. See Sturmbauer (1998) for alternative figure and discussion. See text for details.

absolute space limitations for trophic and reproductive activities would suggest episodes of increased competition in such shrinking environments. Depending upon the extent of crowding, opportunities for hybridization may be extensive (see Mechanisms Driving Differentiation below). With lake level increases, new habitats would become available; during this phase, opportunities for speciation might increase due to rapid population growth and reduced competition. During periods of relatively stable lake levels, speciation and extinction might be at equilibrium or, depending upon the mechanisms driving speciation (below), there might be an increase in species diversity. Overall, species diversity might remain constant at extended high water stands if extinction during regressions balanced speciation during expansion (Fig. 1.2a). However, net species diversity might expand with increases in lake level cycles if new trophically differentiated lineages regularly arose (Fig. 1.2b). With a large number of cycles, if selection during low stands were sufficiently intense, net species diversity might decrease and eventually fluctuate around some equilibrium (Fig. 1.2c). Lake level cycles over long periods of evolutionary time contributed to the differentiation of major lineages in Tanganyika (Sturmbauer 1998), but it is not clear whether such cycles continue to contribute to species diversity.

Sympatry

The origin of tilapias in the Cameroon lakes (Schliwewen et al. 1994) has been extensively acknowledged as demonstrating sympatric speciation (e.g., Page and Holmes 1998), and has been a model for this mode of divergence in cichlids. Analysis of mtDNA sequences of eleven taxa endemic to the Barombi-Mbo crater lake suggested that these species formed a monophyletic group to the exclusion of geographically proximate riverine taxa, presumably the ancestors of this small species flock. Given the characteristics of the species involved, the study suggested that speciation in this endemic fauna was based on trophic differentiation and not sexual selection (Schliwewen et al 1994). In a wide diversity of fishes, radiation via trophic differentiation is a common pattern in recently founded populations (Kruuk 1999). Extremely rapid specialization into planktonic and benthic feeding modes within multiple independent groups suggests the generality of this pattern (Schluter 1998).

Given its centrality, further consideration of the Barombi-Mbo radiation is warranted. Analysis of portions of the molecular data were ambiguous: species trees and gene trees may have been conflated, and sample size limitations may have influenced resolution of these data (Seegers et al. 1999). Potential ancestors from the Niger River, an area that is periodically connected with Cameroon drainages (Roberts 1975), were not examined. As in most phylogenetic investigations, assessment of monophyly depends entirely on the taxa included in the analysis. mtDNA restriction assays of Barombi-Mbo cichlids supported the idea that at least two lineages of tilapias were present in this fauna (S Seyoum & I Kornfield, unpublished information). Regardless, despite the possibility that species flock of Barombi-Mbo may have had more than one progenitor, sympatric speciation appears to have been the mode of origination of much of the current diversity.

A second example of rapid speciation of tilapias has emerged from careful study of the *Alcolapia* (*Oreochromis*) flock of Lakes Natron and Magadi in East Africa (Seegers et al. 1999, Seegers and Tichy 1999). These two small soda lakes support four endemic taxa, three of them syntopic. Like the endemics of Barombi-Mbo, these cichlids are distinguished by morphological features associated with ecological specialization. While speciation in this system may have been sympatric, the complex hydrological history of the area suggests the possibility of repeated lake fragmentation and reunion; analysis of mtDNA sequences is equivocal.

Application of sympatric speciation scenarios to the East African Great Lakes

The extreme youth of the species flocks of the African Great Lakes suggests that some mode of divergence different from the traditionally accepted models of allopatry has operated. It has been suggested that sympatric speciation when it occurs, happens more rapidly than allopatric speciation (McCune and Lovejoy 1998). The inferences of sympatric speciation in the faunas of the Cameroons and in the soda lakes of Natron and Magadi have encouraged workers to consider sympatry as the dominant geographical mode of speciation in the Great Lakes of East Africa (Galis and Metz 1998). Rapid sympatric divergence due to ecological adaptations could apply to initial diversification of lineages in the African great lakes (Rossiter 1995). Under such scenarios, the deepest levels of divergence in any radiation would be based on marked trophic differences. Indeed, conventional (Eccles and Trewavas 1989) and molecular (Sturmbauer 1998) phylogenies of cichlids from both Malawi and Tanganyika are consistent with this idea: major lineages are trophically divergent.

However, models of contemporary sympatric speciation in the African Great Lakes differ from other hypotheses of sympatric speciation. Because there is little evidence that ecological differences are associated with the most closely related species of African haplochromines, sympatric differentiation is presumed to occur by mechanisms not involving rapid ecological divergence, differences in habitat preference or performance mediated by disruptive natural selection. Hypotheses of sympatric speciation in African cichlids facilitated by runaway sexual selection are discussed below.

Mechanisms driving differentiation

Our review of geographical models of speciation in cichlid fishes (above) suggests the validity of various mechanisms by which reductions in gene flow may be realized and wherein evolutionary trajectories may diverge and give rise to new species. We now focus on the forces responsible for the divergence of hundreds of ecologically similar species in Lakes Malawi, Victoria, and Tanganyika. Many of the factors responsible for the great diversity of cichlid species in the African lakes remain unknown. While genetic drift and natural selection are thought to be the primary factors responsible for evolutionary divergence in general, the extent and rapidity of the African radiations have evaded conventional logic in many cases. The most important elements of any divergence scenario are those which lead to testable hypotheses. Table 1.1 summarizes several theoretical frameworks within which cichlid speciation has been studied and emphasizes testable predictions of each model. In this section, we review theoretical and empirical work on speciation by sexual selection following treatment of more conventional evolutionary forces.

Table 1.1 Models of cichlid speciation

Dominant Geographical Model	Mechanisms Driving Speciation	Model System	Testable Assumptions	Author
allopatry	none specified[a]	Malawi *mbuna*	limited gene flow between populations	Fryer (1959)
allopatry	none specified[a]	Victorian haplochromines	biogeography and phylogenetic relationships of extant taxa	Greenwood (1991)
allopatry	(a)sexual selection (b)founder events	Malawi *mbuna*	(a) bottlenecks (b) mate choice on male coloration[b]	Dominey (1984)
allopatry	none specified[a]	Tanganyikan haplochromines	limited gene flow between populations	Rossiter (1995)
allopatry	ruanaway sexual selection on bower size	Malawi sand dwellers	mate choice on bower size/shape[b]	McKaye (1991)
sympatry	runaway sexual selection on male coloration	Victorian haplochromines	(a) mate choice on male coloration[b] (b) genetics of color/preference	Seehausen and VanAlphen (1998,1999)
sympatry	runaway sexual selection	Malawi *mbuna*	(a) mate choice on male coloration (b) genetics of color/preference	Turner and Burrows (1995)

[a] None specified – Mechanisms driving speciation are not emphasized. Classical allopatric divergence sensu Mayr (105) is implied.

[b] Proving mate choice is the minimal first step to inferring speciation by sexual selection. See text for additional consideration.

How important is local adaptation to species diversity?

Some New World cichlids exhibit trophic polymorphisms (Kornfield and Taylor 1983, Sage and Selander 1975) or are phenotypically plastic (Kornfield 1991, Meyer 1987). In the Old World, however, aside from the mollusc crusher *Astatotilapia alluadi* (Greenwood 1965), there are few examples of ecologically significant plasticity. Indeed, there is little evidence to date that these factors contribute to apparent species diversity in the African faunas.

Although cichlids are known for a vast diversity of trophic morphologies, it is generally considered unlikely that differentiation in feeding specializations leads to speciation (Turner 1994). Even in the absence of the necessary phylogenetic information, it is clear that many sibling species differ only in coloration (Ribbink et al. 1983, Turner 1994), while significant trophic morphologies differ at the level of genera (Fryer and Iles 1972, Konings 2001). Unlike the Trinidadian guppy system studied by Endler, which demonstrated natural selection in the form of local adaptation to the presence of predators, many East African cichlids have few natural predators once they are mature. In Lake Tanganyika, predation has been demonstrated to influence reproductive behaviors in some sand dwelling fishes (Karino 1996). Geographical variation in predation pressures has not been documented in Lakes Victoria and Malawi. In general, local natural selection pressures on coloration and behavior, will likely be subtle. Two studies have attempted to correlate coloration with ecological and biological variables and found no (Deutsch 1997) or weak (McElroy et al. 1991) associations. Interspecific competition (Hert 1990) and changes in the biotic community may exert strong pressures in an episodic manner, but their presence and significance has yet to be documented.

Knowledge of local variation in selective regimes would benefit from additional consideration.

Is genetic drift critical to diversification?

Large population sizes are thought to buffer populations against genetic and phenotypic changes, making small population sizes requisite for rapid divergence by drift (Mayr 1963). The postulated effects of random genetic drift following colonization events have been central to several models of cichlid speciation (Dominy 1984, Sturmbauer 1998). Tests of this concept are beginning to appear with the recent advances in molecular genetics; objective protocols for evaluating the magnitude of genetic bottlenecks are available (Luikart et al. 1998). Studies of Malawi mbuna population structure in presumably recently submerged habitat in the southern portions of the lake have been informative. The very low level of mtDNA haplotype variation exhibited *by Melanochromis auratus* (Bowers et al. 1994) is consistent with a historical bottleneck, though current population sizes are quite large. In contrast, an mtDNA study of Malawi zebras showed very little reduction in diversity when narrow endemics were compared with cosmopolitan species (Moran and Kornfield 1995). In such cases, decreases in allelic diversity could also be artifacts of the non-uniform sampling of haplotypes due to small samples sizes and highly variable markers (Ruzzante 1998). Initial surveys of nine Malawi taxa using microsatellite loci showed high levels of allelic diversity (VanOppen et al. 1998).

In two other microsatellite based studies in Malawi, significant population structure was revealed over very small spatial scales, suggesting reduction of gene flow assuming a specific colonization scenario (Arnegard et al. 1999, Markert et al. 1999). These studies correlated loss of observed genetic diversity (either effective number of alleles (Arnegard et al. 1999) or heterozygosity (Markert et al. 1999)) with proxies for isolation (distance from nearest conspecific population (Arnegard et al. 1999) or depth of sand-rock interface (Markert et al. 1999)). Although genetic diversity was consistently high and no evidence for severe bottlenecks was observed, the data suggested a slight loss of genetic diversity after the presumed colonization events. In each case, however, the significance of a bottleneck effect to speciation is speculative. In Lake Victoria, preliminary characterizations of several species using microsatellite markers detected substantial of genetic variation (Wu et al. 1999). Extensive mtDNA- based population studies in the fauna of Lake Tanganyika have revealed considerable variation within a number of narrow endemics (Meyer et al. 1996, Sturmbauer 1998). In summary, there is no general evidence that founder events have contributed to cichlid radiation in the Africa Great Lakes.

What role might hybridization play in faunal evolution?

Hybridization of heterospecific cichlids in nature has been observed to occur as a result of anthropogenic insults or accidental translocations in Lake Malawi (Stauffer et al. 1996, Stauffer and Hert 1992), and by eutrophication of water in Lake Victoria (Seehausen et al. 1997, Witte et al 1997). These observations, and the relative ease of crossing divergent lineages in the laboratory (McElroy and Kornfield 1993) indicate that diversity of sympatric forms in nature is maintained by pre-zygotic, rather than post-zygotic factors. Though distinctiveness is maintained under normal conditions, rapid environmental changes in the past may have contributed to hybridization and the diversity of cichlids observed today. Sudden drops in lake level, for instance, would bring adjacent populations into secondary contact. Color differentiated ecological cognates (for example, previously allopatric conspecifics), might be among the first wave of taxa to hybridize, specifically because of their ecological identity. The results of such hybridizations are unknown and may vary depending on the genetic architecture underlying reproductive traits. Interspecific hybridizations might be facilitated under conditions of extreme population density, perhaps exacerbated by rapid changes in community composition. Whether hybridization can generally cause speciation in animals (Arnold 1997, Dowling and Secor 1997, Grant and Grant 1997) is a subject of speculation; as it is in cichlids (Crapon de Caprona and Fritzsch 1984, Sturmbauer 1998). Certain predictions of the effects of hybridization can be made. In general, levels of genetic variation in hybrids should increase, but distinctive characters such as unique reproductive coloration might be submerged if they are polygenic in nature. Alternatively, if each reproductive character is encoded by separate locus, mosaic phenotypes qualitatively different from parental strains might emerge (Arnold et al. 1999, McElroy and Kornfield 1993). Hybrids that do not suffer a decrease in fitness due to natural selection should thrive, making speciation by hybridization a potentially considerable contributor to overall species diversity. However, if hybrids suffer a decrease in fitness due to natural selection, the evolution of mate discrimination behaviors which facilitate pre-zygotic isolation in areas of secondary contact could be a result of reinforcement – or adaptive increase in mate discrimination (Coyne and Orr 1998). In this case, distinctive male coloration might evolve at the time of secondary contact, rather than during isolation. Reinforcement of female preferences has been modeled in *Drosophila* (Kelly and Noor 1996), and experimentally demonstrated in the three spine stickleback (Rundle and Schluter 1998). The role of reinforcement in cichlid speciation has not been appropriately addressed (Deutsch 1997, Ribbink 1994, Turner 1994) and in part may have been lost in the semantic arguments of species concepts and adaptations (see (Harrison 1998) for a discussion). However, detection of reinforcement in cichlids might be extremely difficult. Variations in numbers of offspring normally observed in species with extremely low fecundities may make detection of postzygotic fitness effects impossible.

Is sexual selection driving cichlid speciation?

That sexual selection operates in cichlid fishes was suggested very early in the study of the group (Kosswig 1947). Although critical empirical support is not abundant, many prominent investigators have assumed in recent reviews that female mate choice on male coloration is a common phenomenon in cichlids (McKaye 1991, Meyer 1993, Ribbink 1994, Sturmbauer 1998, Turner 1994). While interspecific discrimination has been extensively demonstrated (see above for molecular work and (Knight and Turner 1999, Knight et al. 1998, Seehausen and VanAlphen 1998, Seehausen et al. 1997, Seehausen et al. 1998), strong evidence for intraspecific mate choice per se is limited.

Proposed mechanisms of divergence by sexual selection. Models of cichlid speciation by sexual selection fall into two major categories: rapid divergence of sexual characters in allopatry (Dominy 1984, McKaye 1991), and disruptive or divergent sexual selection in sympatry (cichlid refs: Seehausen and VanAlphen 1999, Seehausen and VanAlphen 1998, Seehausen et al. 1998, Turner and Burrows 1995, VanDorn et al. 1998; general refs: Higashi et al. 1999, Kondrashov and Shpak 1998, Lande 1981, Lande 1982). Both scenarios employ the process of Fisherian runaway sexual selection. The distinction of the Fisherian process from other mate choice scenarios, such as 'good genes' or 'handicap' processes, is that the female preference for a sexually selected trait increases in frequency as a result of co-inheritance of trait and preference, rather than via direct selection owing to increased fitness of females who choose quality males (Andersson 1994). Assumptions of the runaway process include (1) heritable genetic variation in male trait(s), (2) heritable genetic variation in female preference, (3) increased mating success of males with exaggerated trait, and (4) genetic covariance of trait and preference (Andersson 1994). Allopatric models posit that geographical variation in sexual characteristics (i.e. coloration and bower form) is generated by founder effects and maintained by limited dispersal. Sympatric models have had to overcome the theoretical restriction imposed by the effects of recombination on reproductive traits and preferences and the elimination of intermediate genotypes (Johnson and Gulberg 1998). This has been done by assuming genetic dominance in female preference traits (Turner and Burrows 1995), variation in habitat fidelity (VanDorn et al. 1998), and extreme mating preferences (Higashi et al. 1999).

Appeal of speciation by sexual selection. The association of species diversity with exaggerated sexual characteristics is indeed appealing. A recent body of literature has addressed this phenomenon among bird species by identifying sexual dimorphism or ornamentation as a correlate of species richness (Barraclough et al. 1995, Moller and

Cuervo 1998, Owens et al. 1999, Price 1998). However, this result begs the question of the causative relationship between these two factors and the necessary elements to determine causation of species richness. Unfortunately, despite the need for rigorous hypothesis testing and the data that are required for empirical support of sexual selection, the ability to infer mechanisms causing species diversity are limited by the time scales in which divergence events occur. Barraclough and his coworkers (1998) suggest several methods for inferring past divergence events using species level phylogenies. Thus, the absolute need for such reconstructions among closely related African cichlid species is clear.

Testing divergence hypotheses. Female cichlids are able to discriminate males of their own species and closely related species by body color (Holzberg 1978, Knight and Turner 1999, Knight et al. 1998, Seehausen and VanAlphen 1999, Seehausen and VanAlphen 1998, Seehausen et al. 1998). This ability is important in the maintenance of diversity, especially where closely related species are sympatric, and not ecologically differentiated (Seehausen et al. 1997). Documentation of interspecific mate choice, though, does not facilitate assessment of competing hypotheses of divergence by sexual selection. Indeed, current data do not allow rejection of a completely allopatric divergence scenario in which sexual selection plays no role. For example, in evaluating evidence of assortative mating of closely related sympatric species (Seehausen et al. 1997), a hypothesis of secondary contact is equally parsimonious with that of divergence in situ.

Intraspecific mate choice has been recorded in field observations (Karino 1997, McKaye 1991) and field experimentation (Hert 1991). Though all of these studies suffer from small sample size and limited replication, they present compelling evidence for the existence of mate choice within species. The male traits under selection in these studies were bower size (McKaye 1991), pelvic fin length and symmetry (Karino 1997), and number of egg spots on the anal fin (Hert 1991). In general, studies have not demonstrated heritability of male traits, female preferences, or the covariation of those traits through breeding experiments.

Alternative hypotheses of the evolution of the extent of sexual dimorphism deserve attention. For instance, intrasexual selection via male contest competition is very likely to operate in cichlids and remains largely unstudied (but see Karino 1996). Divergence of male coloration in allopatry may be as likely under intrasexual as it is under intersexual selection, and intrasexual selection does not require a change in female preference traits. For example, observations previously considered supportive of female choice (McKaye et al. 1993), could be explained by the effects of male competition. McKaye's documentation of female mate choice on bower size and position in Lake Malawi cichlids depends on the use of bower morphology as an 'extended phenotype' of males. Two later studies of mate choice in Lake Tanganyika (Karino 1996) and Lake Malawi (Kornfield et al. 1991) bower building cichlids failed to find evidence for choice in relation to bower morphology, and further found that males are often ousted from their bowers by males who defend territories in the absence of a bower (Karino 1996). These additional data suggest that male competition is important in the acquisition and defense of breeding territiory. Size has been observed to effect competition between males (I Kornfield, personal communication) but the use of conspicuous male coloration as an intermale signal is also an alternative.

In light of the available models (Dominy 1984, McKaye 1991, Seehausen and VanAlphen 1999, Seehausen and VanAlphen 1998, Seehausen et al. 1998, Turner and Burrows 1995, VanDorn et al. 1998) and limited data (Hert 1991, Karino 1997, McKaye 1991, Seehausen et al. 1999), several authors have interpreted their experimental findings as de facto consistent with speciation by sexual selection (e.g., Deutsch 1997). Such associative orientations are of limited value to the evaluation of alternative ideas, since no hypothesis-testing framework is formulated. Future studies will need to address alternative hypotheses more directly so that mechanisms of divergence or components of those mechanisms can be systematically supported or rejected. Phylogenetic hypotheses based on biogeographic patterns and competing methods of divergence are one example of such hypothesis testing (Reithal and Meyer 1997, PF Smith and I Kornfield, in preparation).

Evolutionary generalizations and future studies

In this chapter we have reviewed basic information as well as current research in the study of African cichlid evolution. Discussion of information on cichlids and their natural environments provides the basic data necessary to formulate testable hypotheses regarding the divergence of the many hundreds of cichlid species endemic to the lakes of Africa. For the last several decades, cichlid species flocks (Echelle and Kornfield 1984) have served as model systems in which to study the evolutionary process. As more empirical studies become available to critically evaluate the theories of cichlid evolution, resolution of some issues may be achieved. However, comprehensive understanding of the evolutionary history of these faunas and the processes currently at work remains elusive, and there is growing consensus that a single mechanism may be insufficient to explain the origins of these fishes (McKaye 1991, Sturmbauer 1998). We caution against premature adoption of particular divergence scenarios in the absence of sufficiently strong evidence.

Understanding the genetic basis of relevant traits

Cichlid diversification should be intensively studied at the genomic level. *What is the genetic basis of male phenotypes and of female choice? How is trophic differentiation controlled?* For example, preliminary breeding studies for some Malawi taxa do not support simplistic models for inheritance of coloration (Danley 2001, Kornfield 1991, D McElroy, personal communication). The lack of postzygotic isolation between divergent species lends itself to QTL (quantitative trait loci) analysis of color and behavioral characteristics. The genomic mapping studies pioneered by Kocher and colleagues (1998) will provide a basis for such work.

Understanding coloration in cichlids

The role of color in cichlids is considered central to diversification and is manifest so knowledge about the physiological basis of color production and of color perception is required. *Are cichlids different than other fishes in their use of colors? What is the significance of UV coloration?* For example, in some cichlid systems, limitations to female color perception may be expected to differ because of habitat characteristics, yet sensory ranges may be great. Comparative differences between physiological abilities and abiotic constraints among taxa within and between lakes may reveal processes molding these attributes.

Studying behavior

Male coloration varies extensively, but little is known about variation in female preference. *For what male traits is female preference displayed? Are some types of preferences common to diverse lineages within or between lakes?* Quantification of the magnitude and variation in female preference for coloration (as well as other attributes of male phenotypes) is necessary to understand the process of diversification. Information on search strategies is necessary to better model and to understand the evolution of male traits and female preferences. Importantly, the implications of extensive polyandry and polygyny need also be considered within this framework. Such research is demanding; in situ observations are opportunistic, though territoriality facilitates study of male characteristics subsequent to mating. Experimentation in aquaria must be undertaken with caution since mating is infrequent and artifacts may occur; common proxies for female behavior in choice experiments, such as differential time allocation or proximity to particular males, merit verification. Experimentation under semi-natural conditions may be necessary, e.g., in large ponds where adequate space for territories is present. In some aquatic systems, female choice could be determined by the use of molecular markers with subsequent assessment of male attributes.

Insights from population genetics

Preliminary population studies suggest very limited gene flow among taxa over short distances, but this is contradicted by early studies of migration (McKaye and Gray 1984). *How extensive is geneflow among discontinuously distributed populations. Is isolation by distance a universally demonstrable attribute in these systems?* A test for genetic differentiation at several scales (0.1 to 100 km) in putatively stable populations e.g., those associated with deep basins might provide information on the extent of isolation/gene flow. Observations of population differentiation over geography would suggest that the behavioral characteristics of the species studied are insufficient to produce population subdivision in stable populations in the absence of physical barriers to migration.

Criteria for evaluating models of divergence

Construction of accurate phylogenies is central to evaluate the influence of geography on speciation. *Are molecular phylogenies consistent with hypotheses which can exclude allopatric differentiation?* For example, discrimination between recolonization scenarios, allopatric differentiation, and sympatric divergence can be achieved by examining relationships among clusters of closely related species over large geographic distances (e.g. Reinthal and Meyer 1997). Similarly, with comprehensive geographic coverage, secondary contact can be examined using sensitive molecular frameworks. Species level phylogenies will also provide a basis on which to infer the evolution of important characteristics among sibling species. Such inferences may point to particular characters involved in early divergence.

Chapter 2

PHYLOGEOGRAPHY OF LAKE MALAWI CICHLIDS OF THE GENUS METRIACLIMA: SIGNIFICANCE OF ALLOPATRIC COLOR VARIATION

Abstract

One of the most compelling features of the cichlid fishes of the African Great Lakes is the seemingly endless diversity of male coloration. Color diversification has been implicated as an important factor driving cichlid speciation. Color has also been central to cichlid taxonomy and thus, to our concept of species diversity. Here we undertake a phylogeographic examination of several allopatric populations of the Lake Malawi cichlid, *Metriaclima zebra* in order to reconstruct the evolutionary history of populations, which exhibit one of two dorsal fin colors. We present evidence that populations with red dorsal fins (RT) are not monophyletic. The RT population defining the northern limit of the distribution has evidently originated independently of the southern RT populations, which share a common ancestry among them. This evidence of species level color convergence marks an important discovery in our understanding of cichlid evolution. It implies that divergence in coloration may accompany speciation, and that allopatric populations with similar coloration cannot be assumed to be conspecific. In addition to this finding, we have observed evidence for introgressive hybridization between species, contributing to current evidence that this phenomenon may be extremely widespread. Therefore, in species-level phylogenetic reconstructions consideration must be given the potential effects of introgression. Consequently, our own conclusions must be taken with

the caveat that introgression has affected the genetic diversity among the populations examined, until further studies are available to confirm our result.

Introduction

In Lake Malawi, Africa, the number of endemic cichlid species was recently estimated to be nearly 800 (Snoeks 2000). Much of the species richness in Lake Malawi is in the form of geographic color variation, first described in detail by Ribbink et al. (1983). The importance of male reproductive coloration in the origin and maintenance of new cichlid species in the African Great Lakes is nearly universally accepted. Coloration has been incorporated in models suggesting both sympatric (Seehausen *et al.* 1999, Turner 1995) and allopatric (Danley *et al.* 2000, Dominey 1984) divergence. In spite of the recognized importance of color, little is known about the phylogeographic relationships among allopatric color variants and thus the forces that might drive divergence in coloration.

Taxonomic designation of allopatric color variants is complicated by extreme species diversity and the lack of species-level phylogenies for most groups of cichlids. Allopatric populations with subtle variations in coloration are often assigned to the same species, but the affinities of allopatric populations are not always immediately apparent (Turner *et al.* 2001). When localized color variants cannot be easily allied to more broadly distributed taxa, they are sometimes elevated to specific status. These species usually have distributions limited to the type locality (often a single rock outcropping) and are known as 'narrow endemics' (Ribbink *et al.* 1983). This approach provides a taxonomic framework, but it does not illuminate the ecological and evolutionary events

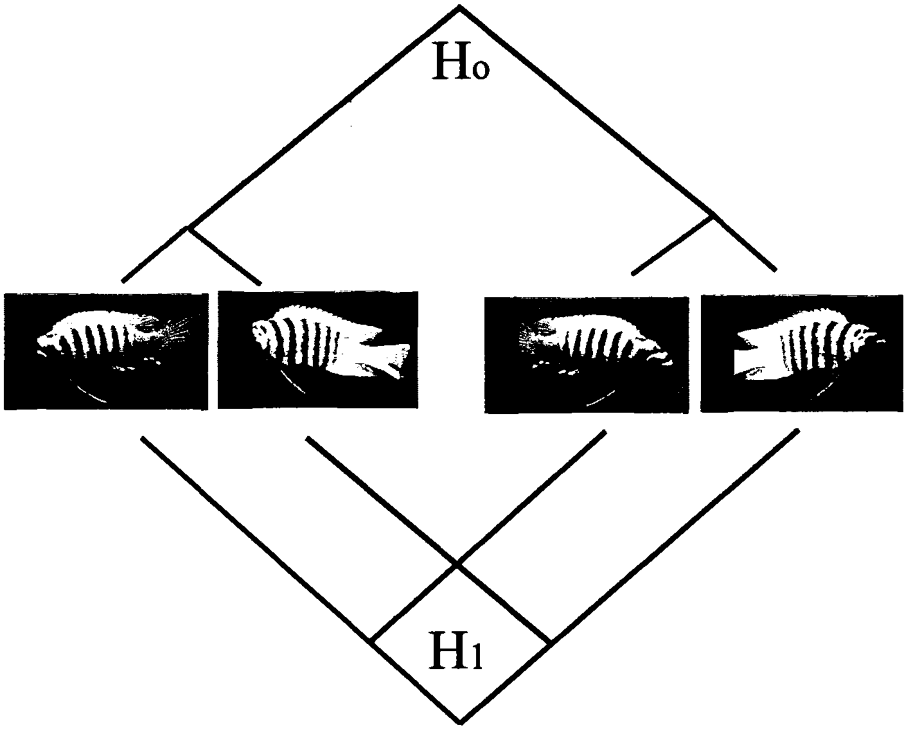

Figure 2.1. Examples of RT an BB phenotypes and graphical representation of the hypotheses tested. H0: Disjunct RT populations arose from respective BB neighbors via convergent evolution. H1: RT populations are monophyletic, and southern distributions are explained by colonization of divergent lineages from the north. The BB pictured is from Thumbi West Island, and the RT is from Nakantenga Island (Photos by Ad Konings).

that have influenced the origin and distribution of cichlid species. Unambiguous groupings of allopatric populations with shared evolutionary histories are necessary for such insights. Because of the difficulty in effectively grouping allopatric variants, and because of the unknown evolutionary stability of coloration, molecular markers are appropriate for generating phylogenetic information. Since ancestral polymorphisms are maintained within the species of the African Cichlidae, DNA sequence variation has not been widely successful in reconstructing ancestral relationships (Moran & Kornfield 1993, Parker & Kornfield 1997), thus we employ microsatellite markers here to investigate phylogenetic relationships.

Climatic variation in Africa's rift valley is associated with dramatic fluctuations in lake levels, a recognized factor that may promote intralacustrine allopatric divergence (Kornfield & Smith 2000). The effect of fluctuations in lake level on cichlid phylogeography has been extensively documented in Lake Tanganyika (Sturmbauer et al. 2001). In Lake Victoria, the origin of the entire species flock of haplochromine cichlids has been estimated at just 14,600 ybp based on inferences from lake level data (Seehausen 2002) (but see Fryer (2001) for an opposing view). In Lake Malawi, one estimate indicated that lake level was lower than 100m below current level just 200–300 ybp, when the southern portion of the Lake would have been completely dry. This implies that many of the southern endemic species and variants have evolved in the dramatically short period since lake level rise (Owen et al. 1990). More recent evidence, however, suggests that the most recent lake level rise was a much older event (10-6 kybp) (Finney et al. 1996).

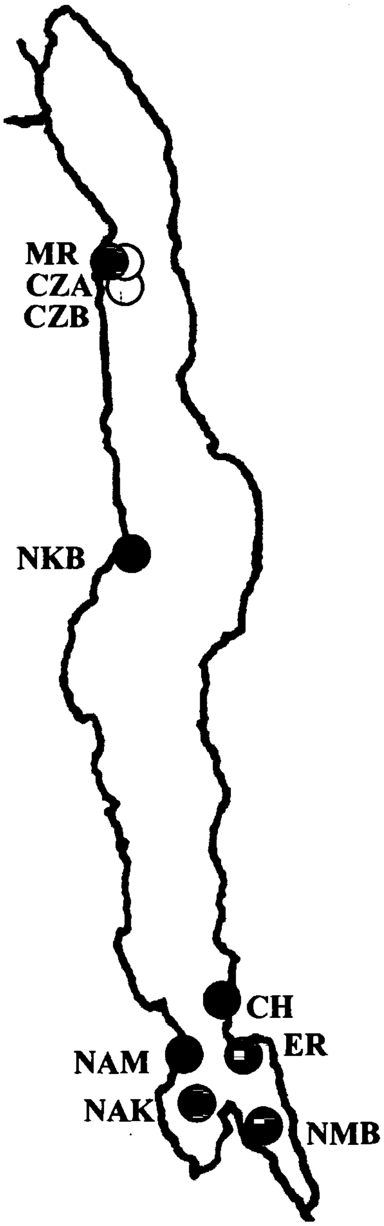

Figure 2.2. Diagram of Lake Malawi showing seven *M. zebra* and two *M. sp.* 'zebra Chilumba' sampling locations. Grey circles indicate locations of RT samples. Black circles indicate locations of BB samples. White circles are the sampling locations of two populations of *M.* sp. 'zebra Chilumba'. Sample codes are used here, sample details are presented in table 1.

Regardless of the exact timing, the depth profile of Lake Malawi is suggestive of a north-to-south colonization trend facilitated by rising water levels at some time in the recent past. Creation of new habitats during lake level rise would certainly have been important in the evolution of the Lake Malawi flock; whether it was accompanied by divergence of new color variants or simply by colonization of northern forms is unknown. Knowledge of the phylogeographic relationships among northern and southern populations and among allopatric color variants is important to developing an understanding of the relative roles of colonization and speciation in the invasion of new habitat by cichlids in Lake Malawi.

Metriaclima zebra is a rock-dwelling Lake Malawi cichlid with lake-wide distribution and considerable geographic color variation; more than twenty species or geographical color variants are currently recognized (Konings 2001). Within the *M. zebra* complex are two well-known forms of zebras, the 'red top' (RT) and 'blue-black' (BB). The RT zebra has a limited and highly disjunct distribution spanning several hundred kilometers from north to south. Four discontinuous populations of the RT zebra have each been recognized as distinct 'narrow endemic' species based on morphometric differences (Stauffer et al. 1997). The BB zebra has a much wider distribution, which is mutually exclusive with the RT zebra. Though these two forms never co-occur (Konings 2001), a hybrid population has recently been reported in one area of the lake (Smith & Kornfield, manuscript in review).

Two explanations could be advanced for the distribution of the RT zebra relative to the BB counterpart: repeated in situ divergence of RT and BB forms in each area of the lake, or colonization of the south by extant, differentiated northern populations (Fig. 2.1).

Similar species distributions in Lake Victoria have been examined as a means of inferring divergence processes. Our null hypothesis is similar to the proposal of Seehausen and van Alphen, that geographically proximate populations within new habitat have originated via (sympatric) speciation (Seehausen & van Alphen 1999). The alternative hypothesis here is aligned with the proposal of Bouton (2000), that such species distributions in new habitat can be explained by colonization of extant lineages. The null hypothesis of in situ divergence, in the present case, requires convergent evolution of the red dorsal fin. Convergence is a phenomenon that is evident across the species flocks of the African great lakes (Kocher et al. 1993), but color convergence at the species level has not been documented. The alternative hypothesis of colonization posits conservation of dorsal fin color and suggests that ecological factors other than lineage splitting (such as habitat preference, competitive exclusion, etc.) are responsible for the current distributions. Because allopatric color variation is abundant in all genera of rock-dwelling cichlids from Lake Malawi (mbuna) resolution of these hypotheses has important implications for our general understanding of mbuna evolution.

We here examine four populations of the RT *M. zebra* and investigate their relationships to each other and to neighboring BB *M. zebra* populations using microsatellite markers. We have analyzed two additional populations of a distinct species in the northern limit of the *M. zebra* distribution (Fig. 2.2) in order to investigate the potential contribution of hybridization to the distribution of genetic variation of *M. zebra*. In addition to resolving these hypotheses, our data may allow generalized inferences regarding questions of north-south biogeography, color pattern evolution, and the potential effects of introgression in the Cichlidae of Lake Malawi.

Materials and methods

Study species

Metriaclima zebra is a widely distributed rock-dwelling cichlid endemic to Lake Malawi. An alternative name, Maylandia (Meyer & Foerster 1984), has been proposed for this genus but is considered a nomen nudem, and *Metriaclima* is used (Stauffer et al. 1997). The rock-dwelling cichlids of Lake Malawi are collectively known as mbuna. Several allopatric color variants of *M. zebra* are recognized and are regarded by some as distinct species. The classic *M. zebra* has a blue dorsal fin and 6-8 vertical black bars over blue ground coloration and we refer to this phenotype as BB. An alternate phenotype exists wherein the dorsal fin is red or orange instead of blue and we refer to this phenotype as RT. BB and RT *M. zebra* never co-occur in the same rock area (Konings 2001, personal observations) and populations are not polymorphic in regards to the RT and BB phenotypes described (Kornfield & Smith 2000, personal observations). Populations of each color form often inhabit adjacent rocky areas, which may be separated by very short distances.

RT populations were sampled from four locations in Lake Malawi: Eccles Reef (*M. thapsinogen*), Nakantenga Island (*M. pyrsonotos*), Mpanga Rocks (*M. emmiltos*), and Namaso Bay (*M. sandaracinos*). Long considered conspecific, these populations were elevated to specific status (species names given in parentheses) by Stauffer et al. (1997). Regardless of taxonomic rank, we seek the mechanism whereby these allopatric populations have differentiated and populated their current habitats. Therefore, to assist the readers' understanding, we loosely refer to all populations examined as variants within the *M. zebra* species complex and note valid species names where appropriate. Populations of BB *M. zebra* were sampled from three locations: Chiofu Bay, Nkhata Bay, and Namalenje Island (Fig. 2.2). Samples of 64 – 72 individuals were collected for DNA analysis at each of these seven locations (Table 2.1).

Table 2.1. Genetic diversity of *M. zebra* populations examined

	Sample Location								
Sample	S Latitude	E Longitude	Code	Phen.	Taxon	N	H_e	H_o	A
Chiofu Bay	13°31.773	34°51.962	CH	BB	*M. zebra*	70	0.842	0.846	22.67
Nkhata Bay	11°36.370	34°17.992	NKB	BB	*M. zebra*	71	0.830	0.836	22.67
Namalenje Is.	13°43.754	34°38.462	NAM	BB	*M. zebra*	64	0.785	0.806	14.56
Eccles Reef	13°45.991	34°57.687	ER	RT	*M. thapsinogen*	72	0.830	0.834	19.89
Nakantenga Is.	13°55.218	34°38.546	NAK	RT	*M. pyrsonotos*	71	0.865	0.856	23.11
Mpanga Rocks	10°25.779	34°16.720	MR	RT	*M. emmiltos*	70	0.832	0.820	21.00
Namaso Bay	14°09.580	35°00.000	NMB	RT	*M. sandaracinos*	67	0.790	0.791	21.11
Mpanga Rocks	10°25.779	34°16.720	CZA	CZ	'zebra Chilumba'	37	0.830	0.765	19.56
Luwino	10°26.775	34°16.966	CZB	CZ	'zebra Chilumba'	55	0.848	0.783	25.00

Phen, Phenotype of population; N, sample size; He, expected heterozygosity (average gene diversity); Ho, observed heterozygosity; A, average number of alleles. See Figure 1 for pictures representing RT and BB phenotypes, and Figure 2 for diagram of sample locations. Long. and Lat. coordinates were calculated using the WGS84 datum.

Two additional *Metriaclima* populations from the Chilumba region in northern Lake Malawi at Mpanga Rocks and Luwino Reef were also examined. These two populations belong to the undescribed *Metriaclima sp.* 'zebra Chilumba' (Ribbink *et al.* 1983) which is thought to be allied with *M. zebra*, but is considered to be outside of the *M. zebra* complex sensu strictu (Konings 2001). This species was not examined to address our proposed hypotheses, rather because it is sympatric with the RT population (*M. emmiltos*) from Mpanga Rocks, and may play a role in introgression.

Sampling & DNA preparation

Fish were collected with monofilament gill nets using SCUBA. Only males were sampled to ensure proper species identification. Small portions of the caudal fin were clipped from individual fish and preserved in 95% ethanol; ethanol was changed after 48 hours. DNA was extracted using the QiAmp tissue kit (Qiagen, Valencia, California) following the manufacturer's instructions.

Nine microsatellite loci were examined: Pzeb2, Pzeb3, Pzeb4, Pzeb5, (Genbank accession numbers X99783, X99784, X99786, X99787 (Van Oppen et al. 1997a)) UME002, UME003, (U14396, U14397 (Parker & Kornfield 1996)), UNH002, UNH231, (U17045, G12382 ((Kocher et al., unpublished)) and DXTUCA3 (U94850 (Seultmann et al., unpublished)). For each sample, alleles at the nine loci were amplified by the polymerase chain reaction (PCR) using standard conditions and separated on an ABI 377 DNA sequencer as per the manufacturer's instructions (Applied Biosystems International). PCR fragments were detected using Genescan (v 3.1) software, and alleles were scored using Genotyper (v 2.1) software (Applied Biosystems International).

Data analysis

Microsatellite data was examined for conformity to Hardy Weinberg expectations and linkage equilibrium using GDA (Lewis & Zaykin 2001). Probability values for Hardy Weinberg were generated to test the fit of genotype proportions to those generated by 3200 random shufflings and were Bonferroni corrected for comparisons across the seven *Metriaclima* populations ($\alpha = 0.05$, critical P-value = 0.0071). An exact test for linkage disequilibrium was used to examine association between alleles among pairs of microsatellite loci. Various combinations of these loci have been used in several other studies (Arnegard et al. 1999, Danley et al. 2000, Markert et al. 1999, Van Oppen et al. 1997b) and have not been shown to be linked. Allele frequencies, heterozygosities and mean number of alleles were calculated using the microsatellite toolkit macro for Microsoft Excel™ (Park 2001). Population pair wise F_{st} values, overall F_{st}, and F_{st} standard deviations were calculated using ARLEQUIN (Schneider et al. 2000).

Divergence between populations was measured with Nei's standard genetic distance (Ds) (1972) and Cavalli-Sforza and Edwards' chord distance (Dc) (1967) using the GENDIST module of the PHYLIP software package (Felsenstein 1993). Van Oppen et al. (2000) found evidence for size homoplasy and non-stepwise mutations for Pzeb4 in Lake Malawi cichlids. This locus (which violates the assumptions of the stepwise mutation model (SMM)) had a great affect on SMM-based distances, but had a much smaller affect on infinite alleles model (IAM)-based distances. Using simulated data, Takezaki and Nei (1996) concluded that Dc was one of the most efficient distances in obtaining correct tree topology with data generated under both the SMM and the IAM. In an empirical study of genetic distances, Paetkau et al. (1997) concluded that Ds was an effective metric for reconstructing relationships between populations that diverged in the last 20 kyrs. Nei's standard distance is based on the IAM, and does not require that microsatellites conform to strict stepwise mutation. The chord distance of Cavalli-Sfvorza and Edwards, which measures the distance between populations in multidimensional space, is free of assumptions of mutational mechanism. Thus, these distances are appropriate metrics for the current data set.

Majority rule neighbor-joining dendograms were generated using the NEIGHBOR and CONSENS modules of PHYLIP (Felsenstein 1993). Data were re-sampled over loci to generate a consensus tree and bootstrap support.

Introgression into the Mpanga Rocks RT population was examined with Factorial Correspondence Analysis (FCA) performed using the program GENETIX (Belkhir 1999). The Mpanga Rocks RT and the *M.* sp. 'zebra Chilumba' samples were compared with the southernmost RT (Namaso Bay) and BB (Namalenje Island) samples.

Results

Microsatellite loci

As noted in previous studies of mbuna populations (Arnegard et al. 1999, Danley et al. 2000, Markert et al. 1999, Van Oppen et al. 1997b), the allelic diversity of microsatellites is very high. In the current study, 11 – 73 alleles were detected per locus with average heterozygosity across populations of 0.452 (Pzeb5) – 0.967 (UNH231) (data not shown). Summary statistics of genetic diversity among populations are provided in Table 2.1. Hardy Weinberg equilibrium was frequently observed among *M. zebra* samples. One locus (Pzeb4) deviated from expected genotype proportions in two populations (NKB & NAM), while another (UME002) deviated from expectations in a single population (CH). The Namalenje (NAM) population had two other departures. All other loci conformed to expectations of Hardy Weinberg equilibrium within all of the *M. zebra* populations examined. The *M.* sp. 'zebra Chilumba' populations were each characterized by multiple Hardy Weinberg deviations. The population at Mpanga Rocks had two loci out of equilibrium (UME002, UME003) and the Luwino population had three (UNH231, Pzeb4, UME002).

Linkage disequilibrium was detected sporadically, with three significant observations (ER, Pzeb2/UME002; NMB, Pzeb2/Pzeb3, Pzeb2/UME003) among all *M. zebra* samples other than Mpanga Rocks (RT), wherein linkage disequilibrium was detected among four of the 36 pair wise comparisons (Pzeb3/Pzeb4, Pzeb3/UME002, UNH231/Pzeb4, Pzeb4/UME003). The genesis of this pattern is not known, but a possible explanation includes introgression from the sympatric *M.* sp. 'zebra Chilumba' population (see discussion). Extensive linkage disequilibrium was detected within both the Mpanga Rocks (15/36 pair wise comparisons) and Luwino (21/36 pair wise comparisons) populations of *M.* sp. 'zebra Chilumba'.

Genetic relationships

Multiple analyses support previous data that populations of *M. zebra* are well differentiated. For all BB and RT samples F_{st} (excluding *M.* sp. 'zebra Chilumba') was 0.055 (s.d. 0.021), while pair wise estimates of F_{st} range from 0.018 – 0.096 (Tab. 2.2). A strong and highly significant correlation between genetic distance and geographic distance was detected among BB and RT samples (Mantel test, $g = 3.54$, $r = 0.801$, $p < 0.01$). The addition of the *M.* sp. 'zebra Chilumba' samples maintains this correlation among all samples in this dataset ($g=5.14$, $r=0.862$, $p<0.01$). This result is highly unexpected because the *M.* sp. 'zebra Chilumba' populations are not believed to be part of the *M. zebra* species complex *sensu strictu* (Konings 2001).

Table 2.2. Pairwise F_{st} (and standard deviations) for *M. zebra* population comparisons.

Populations*		F_{st} Mean†	F_{st} St. Dev.
CH (BB)	ER (RT)	0.046	0.019
CH (BB)	NAK (RT)	0.023	0.010
CH (BB)	MR (RT)	0.023	0.007
CH (BB)	NMB (RT)	0.062	0.03
NKB (BB)	ER (RT)	0.059	0.031
NAM (BB)	ER (RT)	0.091	0.031
NKB (BB)	NAK (RT)	0.037	0.022
NAM (BB)	NAK (RT)	0.072	0.022
NKB (BB)	MR (RT)	0.018	0.004
NKB (BB)	NMB (RT)	0.087	0.055
NAM (BB)	MR (RT)	0.081	0.021
NAM (BB)	NMB (RT)	0.096	0.039
CH (BB)	NKB (BB)	0.025	0.012
CH (BB)	NAM (BB)	0.071	0.026
NKB (BB)	NAM (BB)	0.085	0.030
ER (RT)	NAK (RT)	0.015	0.005
ER (RT)	MR (RT)	0.064	0.026
ER (RT)	NMB (RT)	0.058	0.031
NAK (RT)	MR (RT)	0.040	0.017
NAK (RT)	NMB (RT)	0.044	0.024
MR (RT)	NMB (RT)	0.093	0.052

* see Table 1 for population codes.

†All F_{st} values are significant at the 1% level. For all populations together $F_{st} = 0.055$ (s.d. 0.021).

A neighbor-joining tree was constructed for RT and BB samples from Nei's distance (Fig. 2.3). This tree joined populations that share like color phenotypes with the exception of the Mpanga Rocks (MR) population. The tree shown is the majority-rule consensus of 100 bootstrap replicates; the tree constructed using Dc has the same topology (not shown), but is less resolved. Proportion of replicates supporting each node are shown on the tree. All nodes have bootstrap values of 50 or greater, indicating a well-supported tree topology.

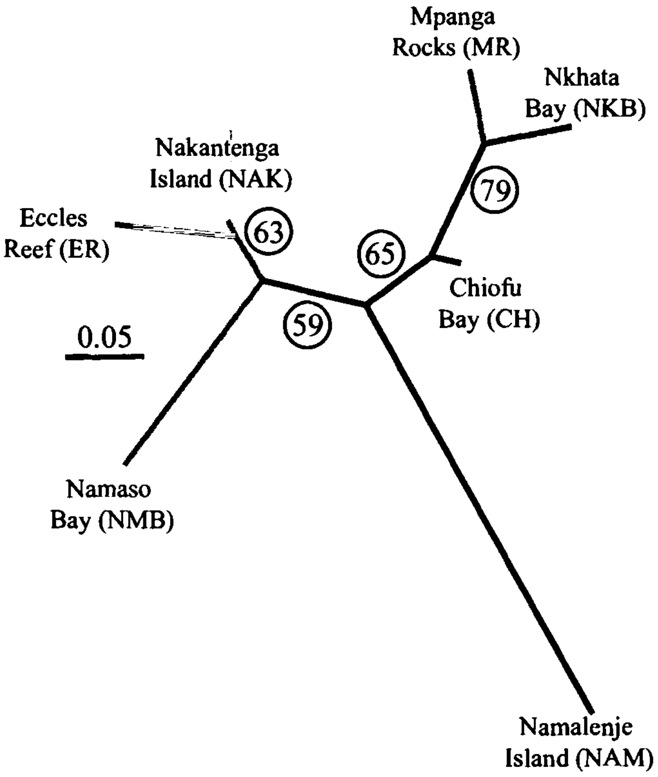

Figure 2.3. Neighbor-joining, majority rule consensus tree of the seven sampled populations using Nei's standard distance (Ds) calculated from nine microsatellite DNA markers. Proportion of one hundred bootstrap replicates (re-sampled over loci) supporting each node are shown in circles. The Mpanga Rocks (MR) RT sample is sister to the northern BB sample from Nkhata Bay, indicating independent origin of the RT phenotype in the north (see Discussion). The tree joins RT samples from the southern portion of the lake, suggesting that this recently available habitat has been invaded by preexisting lineages, rather than by in situ speciation.

The three RT samples that were joined on the neighbor-joining tree are all from the southern portion of the lake, indicating a single origin of the red dorsal fin in the south. The northern RT zebra from Mpanga Rocks (*M. emmiltos*) does not group with southern RT samples, but is sister to the northernmost BB population sample from Nkhata Bay. The results of our phylogeny thus suggest that the RT zebras are non-monopyhletic, and that the red dorsal fin has arisen more than once among *Metriaclima* populations via convergent evolution.

In order to more critically differentiate between ancestral and reticulate relationships we examined the full data set by factorial correspondence analysis (FCA). In a comparison of nine-locus genotypes from northern and southern populations on a FCA plot (Fig. 2.4), the Mpanga Rocks (RT) data are intermediate between the *M.* sp. 'zebra Chilumba' points and the southern RT and BB points but substantially overlap with the *M.* sp. 'zebra Chilumba' cluster. Thus, the RT sample from the Chilumba region (*M. emmiltos*) apparently shares a closer genetic relationship with the sympatric, *M.* sp. 'zebra Chilumba' than with geographically distant RT and BB population samples.

Discussion

This is the first demonstration of convergent evolution of reproductive coloration in cichlids at the species level. The placement of the northern RT (*M. emmiltos*) in the phylogeny supports the null hypothesis that the red dorsal fin has arisen multiple times within *Metriaclima*. However, the three southern RT samples cluster together, suggesting that they share a common ancestor prior to the colonization of their respective habitats; thus, they have not arisen independently within their southern habitats from their BB

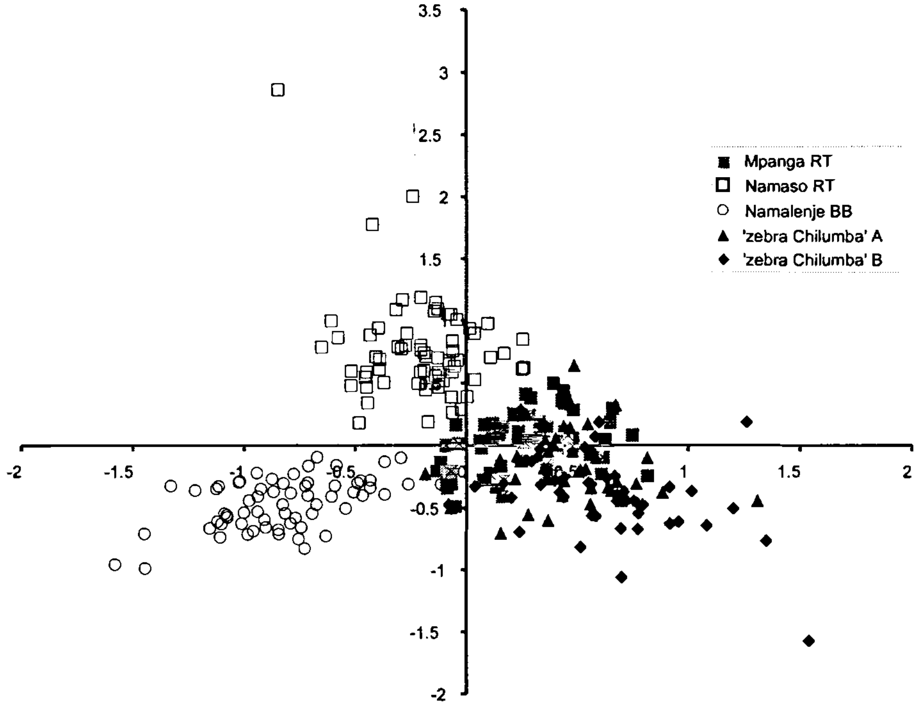

Figure 2.4. Factorial Correspondence Analysis (FCA) shows strong separation of *M.* sp. 'zebra Chilumba' individuals from the southern BB (Namalenje Island) and RT (Namaso Bay) individuals. Mpanga Rocks RT individuals occupy much of the space intermediate between *M.* sp. 'zebra Chilumba' and the southern samples, and substantially overlap the cluster of *M. sp.* 'zebra Chilumba' individuals. The overlap of the Mpanga Rocks individuals with the sympatric *M. sp.* 'zebra Chilumba' individuals and their exclusion from southern *M. zebra* clusters implicates introgressive hybridization in influencing the genetic makeup of the Mpanga Rocks RT population.

neighbors. This result indicates the invasion of newly available southern habitat by distinct lineages that were present in the deeper portion of the Lake prior to lake level rise (Bouton 2000).

Convergent evolution of reproductive coloration is an important finding for our understanding of mbuna evolution. Allopatric populations with similar coloration have long been considered conspecific (Ribbink et al. 1983) and differences in coloration considered sufficient to delimit species (Stauffer et al. 1997). The data presented here indicate that both of these assumptions may be untrue, as RT populations appear paraphyletic, and the populations joined with the greatest support differ in coloration. Color convergence is prevalent at the generic level. For example, species in other Malawi mbuna genera (eg. *Cyanotilapia afra* & *Labeotropheus fullerborni*) have a similar body coloration to the RT examined here (Konings 2001) and have evolved independently of each other (Albertson et al. 1999). Our data indicate that coloration may be genetically labile, and convergence may be ubiquitous even at the species level.

The role of coloration in species divergence remains unknown, but sexual selection has been proposed as a mechanism that might drive speciation (Dominy 1984, Seehausen & van Alphen 1999, Seehausen et al. 1999, Turner 1995, Turner 2000). Sibling species that diverged by sexual selection are expected to differ in selected characters of coloration. The species level color convergence apparent in our phylogeny indicates that change in coloration may have a role in the process of speciation, and that dorsal fin color holds potential interest for study as a sexually selected character. This finding suggests that speciation, marked by divergence in male coloration, may be extremely rapid and that many cryptic species bearing similar color but derived of independent origins may exist.

The phylogenetic signal observed in this dataset may be substantially influenced by introgression. The RT population at Mpanga Rocks (*M. emmiltos*) is sympatric with a *M.* sp. 'zebra Chilumba', a species which is morphologically and behaviorally distinct but bears coloration similar to the *M. zebra* populations examined here (Ribbink et al. 1983). We detected linkage disequilibrium among loci within the Mpanga Rocks RT population and extensively within both *M.* sp. 'zebra Chilumba' populations. Linkage disequlibrium is characteristic of populations experiencing introgression (Hedrick 1999). Examination of the two populations of *M.* sp. 'zebra Chilumba' indeed reveals remarkable genetic similarity to the sympatric Mpanga Rocks RT population. Additionally, the correlation between geographic and genetic distance among all populations is maintained with the addition of these populations. Isolation by distance among several allopatric populations, many of which believed to represent good biological species, is highly unexpected and implies a role for geneflow across species boundaries in determining the partitioning of genetic variation among populations of Lake Malawi cichlids. This phenomenon certainly deserves additional examination in future studies. Importantly for the current study, these data along with FCA plot suggest that the genetic variation in the Mpanga Rocks RT population may have been influenced by introgression from the sympatric *M.* sp. 'zebra Chilumba'.

The observation of introgression in the northern RT population does not invalidate the independent RT origin indicated by our microsatellite phylogeny. However, introgression might explain the sister relationship between the Nkhata Bay BB and the Mpanga Rocks RT. If a cline of gene frequencies existed between Chilumba and Nkhata Bay at the time of colonization of the northern RT, introgression might produce the phylogenetic signal observed here. Such a cline should be observable in the *Metriaclima* populations along the coast between Chilumba and Nkhata Bay as a test of this hypothesis. Genetic complementation tests have been performed with northern and southern RT populations, and are consistent with a common genetic origin for the RT phenotype (Reneau 2001), which may be at odds with our phylogeny.

Geneflow between the two well differentiated species at Chilumba (*M.* sp. 'zebra Chilumba and *M. emmiltos*) suggests that selection may be operating to maintain distinct color phenotypes within both taxa. In depth examination of the color phenotypes at each locality in the Chilumba region and field and laboratory examination of mate choice behaviors among these taxa would be highly relevant to further assessing the nature of this potential hybrid zone. Additional phylogeographic data should help to resolve these relationships further, and will be necessary before we fully embrace the hypothesis of multiple origins of the RT phenotype within the *M. zebra* species complex.

The putative hybridization evident in this phylogeny, together with recently published accounts (Ruber et al. 2001, Salzburger et al. 2002), suggests that this phenomenon may be much more widespread than previously proposed. If secondary contact often results in geneflow, attempts at species-level phylogenetic reconstruction will be challenged with unraveling the signal of historical relationships from genetic similarities generated by reticulation.

This study is the first to use microsatellites to reconstruct a species level phylogeny for cichlids, and shows the effectiveness of the methodology for this group.

Though the microsatellite phylogeny produced here is robust, evidence for extensive introgression within the most critical branch of our tree demands that further examination be directed at the problems of phylogeny and color convergences at the species level in Lake Malawi cichlids.

Chapter 3

HYBRID ORIGIN OF A CICHLID POPULATION IN LAKE MALAWI: IMPLICATIONS FOR SPECIES DIVERSITY AND GENETIC VARIATION

Abstract

Lake Malawi is home to an extremely diverse endemic fauna of cichlid fishes. Both the diversity of species within the Lake and the genetic variation within species have yet to be explained. However, the importance of species recognition to the generation and maintenance of diversity has been repeatedly stressed. Hybridization – the apparent breakdown of species recognition - has been observed sporadically among African cichlids and has been viewed as both a constructive and a destructive force with respect to species diversity. Here we describe a hybrid population in the *Metriaclima zebra* complex within Lake Malawi. This population has a unique color phenotype, high levels of genetic variation, and has apparently stopped exchanging migrants with the putative parent populations. We provide genetic evidence of the hybrid origin of this population, and discuss the potential evolutionary consequences of interspecific hybridization in Lake Malawi cichlids. We propose that hybridization may be an important evolutionary phenomenon with explanatory power concerning both species diversity and genetic variability of Lake Malawi cichlids.

Introduction

Several hundred species of endemic cichlids have arisen in Lake Malawi since its origin less than 1mya. The mechanisms of the evolution and maintenance of this species diversity are still largely unknown (Kornfield & Smith 2000). Ecological differentiation may be central to the long-term coexistence of divergent species, but prezygotic reproductive isolation is critical in the early stages of species divergence among African cichlids (Danley & Kocher 2001). Sexual selection on male color has been proposed as a mechanism that generates cichlid species diversity by causing speciation in sympatry (McKaye 1991, McKaye et al. 1993, Seehausen 1999, Seehausen & van Alphen 1999, Seehausen et al. 1999) or allopatry (Dominey 1984, Danley et al. 2000), and coloration is certainly important to maintain separation between divergent cichlid populations after secondary contact (Seehausen et al. 1998). Indeed, mate choice has been demonstrated as a means of species recognition and species specific mating in cichlids (Knight et al. 1998, Knight & Turner 1999). A corollary to the accepted importance of mate choice is that the lack of species specific mating, or hybridization, could also have important evolutionary consequences that may be central to our understanding cichlid diversification.

The evolutionary outcomes of hybridization are unpredictable; from a parochial perspective, hybridization destroys species barriers and threatens species diversity. However, if hybridization results in novel phenotypes and viable offspring, while leaving parental species intact, it could increase species diversity by generating a hybrid taxon. Regardless of its effect on species diversity, when hybridization results in introgression, the genetic variation within populations will be increased. With modern molecular biological techniques, the detection of hybridization has become possible, and the potential roles of hybridization as an evolutionary force are becoming better appreciated (Allendorf et al. 2001).

Phylogenetic studies of African cichlids have revealed a great deal of genetic variation within species. The maintenance of ancestral mtDNA polymorphism due to incomplete lineage sorting (Moran & Kornfield 1993, Parker & Kornfield 1997), and the "excess" diversity observed at microsatellite loci (Van Oppen et al. 1997b, Markert et al. 2001) have confounded attempts to reconstruct the relationships among Malawi cichlids. These data are difficult to reconcile with the known history of Lake Malawi (Johnson 1996) and its biota (Meyer 1993). Using coalescent methods to trace extant alleles present at mtDNA markers, extremely large effective population sizes have been inferred to explain the maintenance of such genetic diversity over the geologically limited age of the species. Intergeneric hybridization was advanced as an alternative explanation for shared polymorphism among divergent genera. This explanation was viewed as unlikely because of the extent of hybridization that would be necessary, and the apparent absence of intergeneric hybrids in nature (Parker & Kornfield 1997). Recent phylogenetic data have further confirmed that intergeneric hybridization is not likely to be a general phenomenon (Albertson et al. 1999). However, among congeners, if secondary contact of distinct allopatric populations or newly derived species often results in geneflow, genetic effective population sizes will be much higher than expected from census sizes. Thus, the problem of excess polymorphism might be explained by hybridization among closely related species rather than between genera.

Hybrids have been observed in nature among the cichlids of the three African Great Lakes. Putative hybrids between distinct species were discovered in Lake Victoria, where species specific mating was compromised by reduced visibility resulting from eutrophication caused by human activities (Seehausen et al. 1997, Verschuren et al.

2002). In Lake Tanganyika, introgression is believed to have occurred upon secondary contact between a pair of sympatric species (Ruber et al. 2001), and has caused speciation in a separate Lake Tanganyika event (Salzburger et al. 2002) In Lake Malawi, some specimens of intergeneric hybrids have been discovered at the site of an intralacustrine translocation (Stauffer et al. 1996), and hybridization was also implicated in the morphometric differentiation of translocated populations of *Metriaclima callainos* from their source populations (Stauffer & Hert 1992). Additional recognition of hybridization, especially in undisturbed settings, is important to evaluating the potential for hybridization as an evolutionary mechanism working in African cichlids.

Within the *Metriaclima zebra* species complex in Lake Malawi there are multiple populations that differ in dorsal fin coloration. 'Red top' (RT) and 'blue-black' (BB) forms have a mutually exclusive, checkerboard distribution, and uniform dorsal fin phenotype within each population (A. Konings, I. Kornfield and P. Smith, personal observations). *Metriaclima* has been the subject of a taxonomic revision that saw the elevation of four RT populations to specific status, and retention of a single nomen for all BB populations (bearing the blue dorsal fin) (Stauffer et al. 1997). Areas of suitable *M. zebra* habitat are effectively islands; sandy bottom and deep water separating rocky areas may be strong barriers to dispersal (Markert et al. 1999). No obvious zones of contact or clines have previously been observed between the two color forms, even when they inhabit adjacent rocky habitat. Because of the contemporary absence of sympatry, prezygotic reproductive isolation among allopatric variants of the *M. zebra* complex has not been tested in nature, but aquarium studies reveal no postzygotic barriers (Kornfield

1991). Thus, periods of sympatry could lead to hybridization between RT and BB populations .

During fieldwork in 1998, we encountered a population of *Metriaclima* on the eastern shoreline of Lake Malawi exhibiting a mosaic of phenotypes intermediate between RT and BB. This population, located at Makanjila, is situated midway between the RT population at Eccles Reef and the BB population at Chiofu Bay (Fig. 3.1a, 3.1b). The RT at Eccles Reef is assigned to the nominal taxon, *Metriaclima thapsinogen*, and is thus specifically distinct from the *M. zebra* at Chiofu. Because these two species exist only in allopatry, nothing is known about reproductive isolation, or potential isolating mechanisms if they were to become sympatric; maintenance of distinct phenotypes indicates the absence of geneflow between these populations. Several other species inhabit the reef at Makanjila, which is shallow compared to the two adjacent localities. The substrate consists of rock clusters scattered on a sandy bottom, as opposed to the completely rocky habitats of Eccles Reef and Chiofu Bay. Territorial males in the Makanjila population exhibited a range of dorsal fin color phenotypes, but none could be characterized as the typical RT or BB. The intermediate coloration of this population and its geographic location between the *M. thapsinogen* population at Eccles Reef and the *M. zebra* at Chiofu Bay are suggestive of a hybrid origin.

In this study we use microsatellite markers to infer the history of the phenotypically intermediate Makanjila *Metriaclima* population, and examine its relationship to the geographically flanking RT (*M. thapsinogen*) and BB(*M. zebra*) populations. In the absence of diagnosable genetic characteristics, determination of hybrid identity is problematic. We compare the Makanjila population to the parental

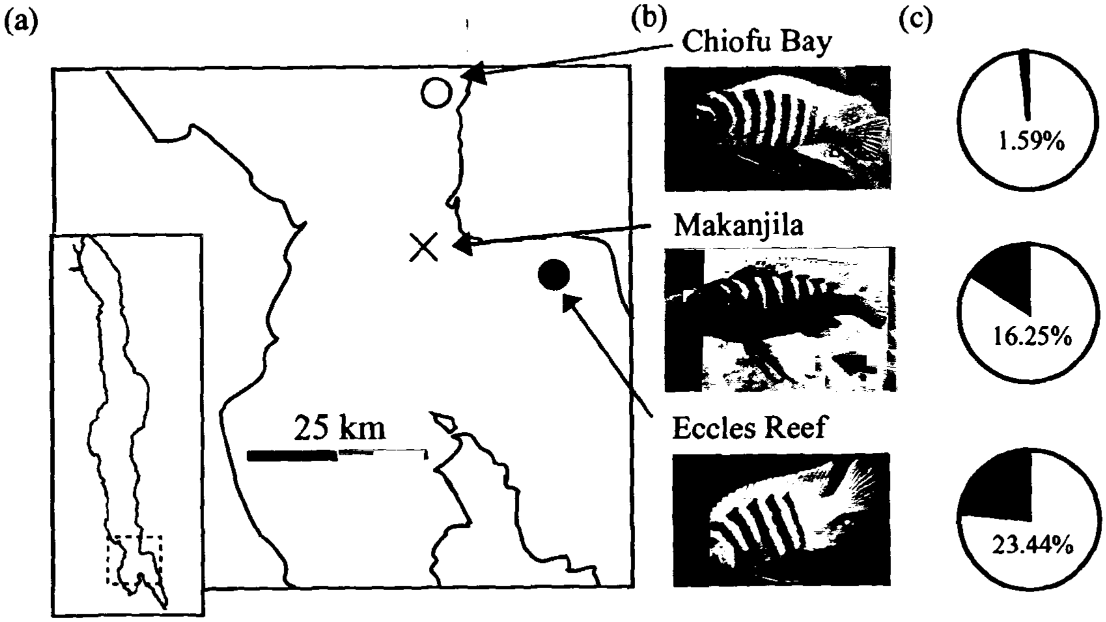

Figure 3.1. A hybrid population in southeastern Lake Malawi. (a) The hybrid population is located at Makanjila, approximately midway between the putative parental populations. (b) The phenotype of the Makanjila hybrid population is intermediate to the two putative parental populations. (c) Frequency of alleles in the saltine distribution of microsatellite locus UME003 (black portion of pie charts) is intermediate in the Makanjila population.

populations and to a virtual population, simulated as F_1 hybrids between parental populations. This simulated hybrid sample provides a benchmark for our analyses.

Materials and methods

Fish were collected with monofilament gill nets using SCUBA. Only males were sampled to ensure unambiguous species identification. Small portions of the caudal fin were clipped from individual fish and preserved in 95% ethanol; ethanol was changed after 48 hours. DNA was extracted using the QiAmp tissue kit (Qiagen, Valencia, CA.) following the manufacturer's instructions.

Four microsatellite loci were examined: Pzeb3, Pzeb4, Pzeb5, (Genbank accession numbers X99784, X99786, X99787 (Van Oppen et al. 1997a)) and UME003 (U14397 (Parker & Kornfield 1996)). For each sample, alleles at the four loci were amplified by the polymerase chain reaction (PCR) using standard conditions and separated on an ABI 377 DNA sequencer as per the manufacturer's instructions (Applied Biosystems International). PCR fragments were detected using Genescan (v 3.1) software, and alleles were scored using Genotyper (v 2.1) software (Applied Biosystems International).

In the absence of fixed diagnostic alleles in each parental population, genetic relationships were examined via comparison to a simulated hybrid population. The simulated F_1 hybrid population (SIM) was generated by combining alleles randomly drawn from the frequency distribution of each putative parental population to create each individual. This process was repeated with each locus to create a sample of 50 F_1 multilocus genotypes. This simulation models the generation of an F_1 hybrid population by extensive hybridization between the putative parental populations. Though allele frequencies might change during successive generations within a true hybrid population, the SIM sample provides a simple snapshot of allele frequency expectations for an F_1 population created by extensive hybridization. The SIM population was subsequently treated identically to the other populations during analysis.

Tests of Hardy Weinberg equilibrium, linkage equilibrium, and population differentiation were performed with Genepop (Raymond & Rousset 1995). Allele frequencies, heterozygosities and mean number of alleles were calculated using the microsatellite toolkit macro for Microsoft Excel™ (Park 2001). Assignment tests were performed with WHICHRUN (Banks & Eichert 1999). Factorial Correspondence Analysis was performed on the data to examine genetic structuring among the three populations. The four locus genotypes of each individual were used as variables and FCA was performed as implemented in the computer program GENETIX (Belkhir 1999)

Sequencing of microsatellite alleles was performed by cloning PCR products from individuals typed for the allele of interest. PCR cloning was done with the TA cloning kit (Qiagen, Valencia, CA); colonies were screened for inserts using blue-white screening on ampicilin impregnated plates. Cycle sequencing was performed using the Big Dye terminator kit (Applied Biosystems International) followed by electrophoresis on an ABI 373 (Applied Biosystems International).

Results

The Makanjila population has enhanced genetic diversity (Nei's unbiased gene diversity) compared to the Eccles Reef *M. thapsinogen* and Chiofu Bay *M. zebra* populations, but is similar to the simulated F_1 hybrid (SIM) population. However, a significant overall heterozygote deficiency was observed in the Makanjila population as well as a positive inbreeding coefficient (Tab. 3.1). Each population is phenotypically distinct and the samples of putative parental populations were more than 25km apart, with no parental-types observed at the Makanjila location. This physical separation and phenotypic distinctivness eliminates the possibility that the heterozygote deficiency is due to a Whalund effect caused by pooled sampling of distinct gene pools. It is likely that inbreeding has occurred within the Makanjila population, despite high levels of genetic diversity.

Table 3.1. Genetic diversity of populations examined

Population	n	H_e	H_o	F_{is}	A
M. zebra	63	0.6957	0.6984	-0.0040	11.75
M. thapsinogen	64	0.6815	0.7031	-0.0310	11.50
Makanjila	40	0.7269	0.6562*	0.0980	11.75
SIM	50	0.7246	0.7250	-0.0004	12.25

(n) number of samples, (He) Nei's unbiased gene diversity, (Ho) observed heterozygosity, (Fis) method of moments estimate of inbreeding coefficient, (A) average number of alleles.

*Significant global heterozygote deficiency ($p = 0.0095$)

A microsatellite saltine distribution (Kornfield & Parker 1997) was identified at locus UME003 in a lake-wide population survey of RT and BB *Metriaclima* populations (P. Smith & I. Kornfield, unpublished data). A large proportion of the alleles observed in RT populations throughout the lake occurred in the lower portion of this disjunct allele frequency distribution (allele size < 177bp), while BB populations had low frequencies of these alleles. The cumulative frequency of alleles within this saltine distribution in the Makanjila population is intermediate between the high frequency (>23%) in the Eccles

Reef (*M. thapsinogen*) population and the minimal (<2%) frequency in the Chiofu Bay (*M. zebra*) population (Fig. 3.1c). This intermediacy is suggestive of a hybrid origin. Indeed, examination of the allele frequency distributions for the other loci (Tab. 3.2) reveals intermediate frequencies for the Makanjila population (and similarity to SIM); this is most evident for alleles with relatively high frequency in one of the putative parental populations. Thus, the allele frequencies generally fit the pattern consistent with the creation of a hybrid taxon via an episode of extensive hybridization between RT *M. thapsinogen* and BB *M. zebra* populations. However, there are exceptions to this pattern, which hint at a more complex history for the Makanjila population.

For three of the loci in this study (Pzeb3, Pzeb4, and UME003) a total of twelve alleles were detected only in the Makanjila population. However, some of these alleles have very low frequencies, and may not be unique. To determine the significance of unique alleles, we performed a simulation of population sampling as an evaluation of our empirical allele frequency distribution. For its frequency to be considered "significantly unique", an allele must be sampled with greater than 95% confidence in our simulation. Table 3.2 presents the allele frequencies for the three sampled populations (and the SIM population) for four microsatellite loci. Four alleles were significantly unique to the Makanjila population based on our simulation. If these alleles exist in equal frequencies in the flanking populations, they would be sampled with > 95% confidence.

Table 3.2. Allele frequencies (as percent) for three populations and the simulated hybrid at four microsatellite loci.

Locus	Allele	*M. zebra*	*M. thapsinogen*	Makanjila	SIM	Locus	Allele	*M. zebra*	*M. thapsinogen*	Makanjila	SIM
PZeb3	310*	...	...	2.50	...	UME003	111	...	0.78	...	1.00
	312	44.44	64.06	52.50	58.00		113*	...	...	1.25	...
	314	3.17	13.28	10.00	5.00		115*	...	...	2.50	...
	316**	...	...	5.00	...		117	...	1.56		...
	318	26.19	...	21.25	10.00		121	...	0.78	2.50	...
	320	8.73	21.09	1.25	16.00		123	...	1.56	...	5.00
	322	1.59	0.78	1.25	2.00		125*	...	...	2.50	...
	324	6.35	0.78	3.75	4.00		127	...	4.69	...	4.00
	326	2.38	...	1.25	4.00		129	...	3.13	...	...
	328	0.79	...	...	...		131	0.79	1.56	1.25	1.00
	330	1.59	...	...	...		133	0.79	5.47	...	4.00
	332	3.17	...	...	1.00		135	...	1.56	...	...
	336	1.59	...	...	...		141	...	0.78	...	2.00
	340*	...	...	1.25	...		143*	...	...	3.75	...
							147	...	0.78	...	...
PZeb4	112	3.97	...	...	4.00		149	...	0.78	1.25	...
	120	36.51	28.13	30.00	34.00		151*	...	...	1.25	...
	124	9.52	44.53	16.25	21.00		177*	...	...	1.25	...
	126**	...	...	32.50	...		181	...	3.13	...	3.00
	128	34.92	3.13	18.75	21.00		185	...	0.78	...	...
	130	3.97	...	1.25	2.00		189	2.38	...	6.25	...
	132	0.79	...	...	...		191**	...	...	12.50	...
	134	0.79	0.78	1.25	2.00		193	...	1.56	2.50	1.00
	136	6.35	21.09	...	10.00		195**	...		7.50	...
	138	3.17	...	...	3.00		197	3.17	16.41	1.25	7.00
	148	...	2.34	...	3.00		199	...	0.78	2.50	...
							201	...	3.13	6.25	1.00
PZeb5	123	5.56	40.63	28.75	29.00		203	0.79	3.91	...	5.00
	129	76.98	51.56	62.50	63.00		205	0.79	1.56	6.25	1.00
	137	17.46	7.81	8.75	8.00		207	...	3.13	7.50	1.00
							209	1.59	0.78	5.00	1.00
							211	3.17	3.91	3.75	5.00
							213	7.94	7.81	3.75	7.00
							215	4.76	7.81	2.50	5.00
							217	6.35	4.69	...	5.00
							219	7.14	2.34	...	2.00
							221	8.73	5.47	3.75	7.00
							223	4.76	3.91	1.25	5.00
							225	6.35	3.91	2.50	6.00
							227	7.14	...	3.75	3.00
							229	9.52	0.78	...	6.00
							231	6.35	...	1.25	4.00
							233	5.56	...	2.50	3.00
							235	3.97	...	...	2.00
							237	3.17	0.78	...	2.00
							239	2.38	...	...	...
							241	1.59	...	...	...
							243	0.79	...	...	1.00

* Allele unique to Makangila population but not significant based on simulation (see text).

** Allele unique to Makangila and significant based on simulation (see text).

In order to test whether alleles may have arisen as new mutations, we sequenced allele '126' from locus Pzeb4. This allele, which has a frequency of over 30% in the Makanjila population, is absent from the other two samples. Locus Pzeb4 was previously examined in-depth for the presence of convergent allele sizes in Lake Malawi cichlid populations. That study revealed that alleles with the same gel mobility often have different sequences. If allele '126' originated by mutation in the Makanjila population, it may have a different sequence than the same sized allele in other *Metriaclima* populations. Alleles in the '126' size class were sequenced from the Makanjila population as well as the *M. zebra* (BB) populations from Nkhata Bay and the Namalenji Island, two distant localities representing northern and southern portions of the *M. zebra* range, respectively. All '126' alleles sequenced were identical. The sequence obtained matched the sequence reported by VanOppen et al. (2000) as allele '129'. The nominal allele size difference between the two studies is a result of the different techniques used, and is not relevant to assessment of homology. Our data do not indicate that this allele arose by mutation in the Makanjila population, but this possibility cannot be definitively ruled out.

Examination of the relationships among populations reveals a general trend of intermediacy of the Makanjila population which is similar to the SIM population. An exact test of population differentiation (Tab. 3.3) is significant for Eccles Reef *M. thapsinogen* and Chiofu Bay *M. zebra* populations at all four loci, with less complete (fewer loci significant) differentiation among all other population comparisons. A genotype assignment test of individuals to the *M. thapsinogen* and *M. zebra* baseline populations (Fig 3.2) shows accurate assignment of individuals from each of those populations with intermediate assignment values of the SIM samples, which displayed non-preferential assignment between the two tested baseline population samples. Makanjila samples are also assigned non-preferentially, but genotype probability values are substantially lower for individuals from this population in both baselines (points fall farther from the origin on both axes). These very low values may be caused by the presence of unique alleles (Tab. 3.2, discussed below) in this population.

Table 3.3. Exact tests of population differentiation by locus

Population Comparison		PZeb3	PZeb4	PZeb5	UME003
M. zebra	*M. thapsinogen*	**	**	**	**
M. zebra	Makangila	*	**	**	**
M. zebra	SIM	N.S.	N.S.	**	N.S.
M. thapsinogen	Makangila	**	**	N.S.	**
M. thapsinogen	SIM	**	**	N.S.	N.S.
Makangila	SIM	**	**	N.S.	**

p-values are shown. Significance values are Bonferroni corrected.
* $p < 0.05$; ** $p < 0.01$; N.S. - not significant

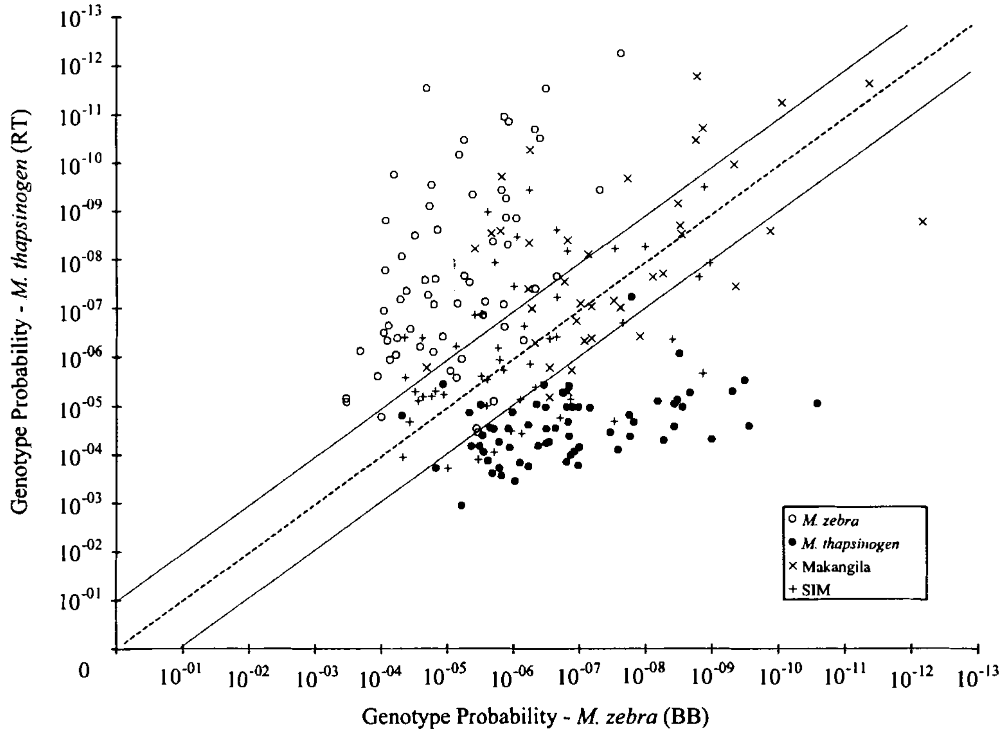

Figure 3.2. Genotype assignment test (four loci) of all samples to putative parental baselines (M. thapsinogen and M. zebra). Individuals were assigned a genotype probability for each baseline and plotted on a log scale. The dashed line indicates the region of the graph where the probability of a genotype in the two baseline samples is equal; samples falling above the line are 'assigned' to the M. zebra (Chiofu Bay) population and samples below the line are 'assigned' to the M. thapsinogen (Eccles Reef) population. The solid lines deviate from equality by a factor of ten. M. thapsinogen and M. zebra samples assigned to their own baseline with a few minor deviations. SIM samples assigned indiscriminately to the two baselines, as did the Makanjila samples; thus these points form clouds around the center of the graph, rather than skewed to either baseline. Some Makanjila samples had very low genotype probability values (points farthest from the origin on both axes), likely caused by unique alleles in that population.

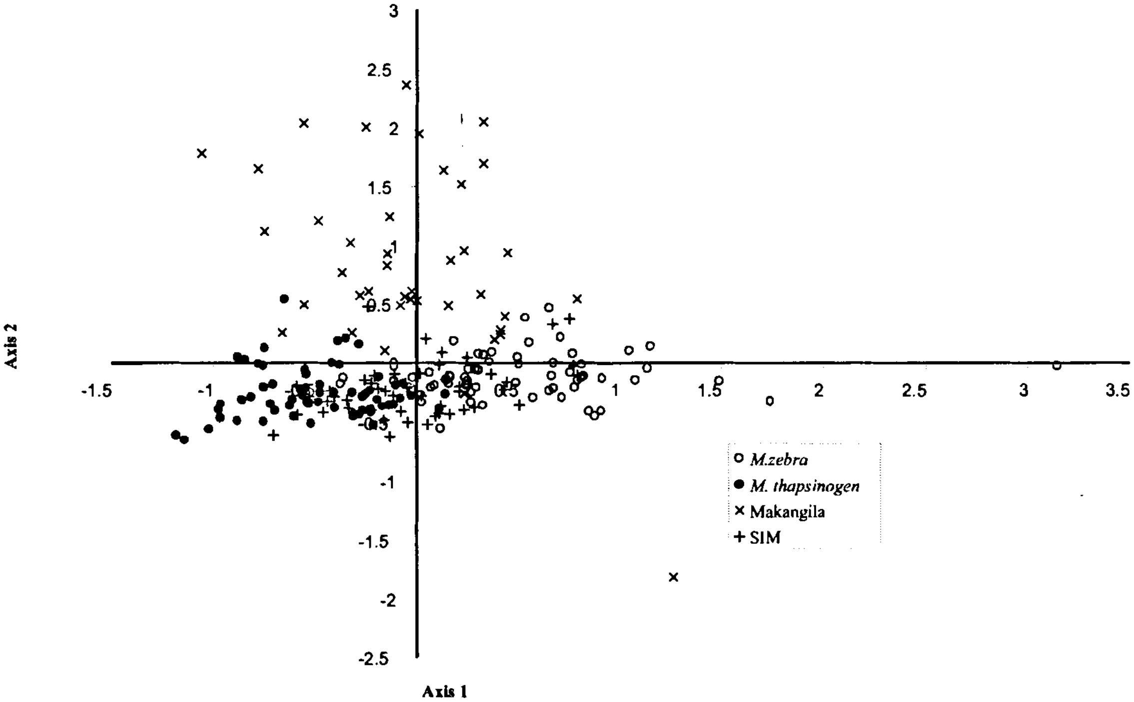

Figure 3.3. Graphical representation of genetic variance among samples by factorial correspondence analysis (FCA). The first two axes are shown; additional axes provide no additional separation among these groups. Axis 1 explains 3.46% of the variance and clearly separates the M. thapsinogen (Eccles Reef) and M. zebra (Chiofu Bay) samples. The SIM and Makanjila samples cluster more centrally, and are completely contained within the range set by the parental taxa on this axis. The 2nd axis explains an additional 3.37% of the variance, and separates some Makanjila individuals from all others in all populations. The alleles predominantly contributing to axis 2 are all unique to the Makanjila population; removal of samples with unique alleles from the analysis results in the remaining Makanjila samples being indistinguishable by axis 2 (data not shown).

Factorial correspondence analysis (FCA) was used to graphically represent the partitioning of genetic variance among populations (Fig. 3.3). Again, clusters of samples from the Eccles Reef *M. thapsinogen* and Chiofu Bay *M. zebra* populations are well separated by axis 1 of the FCA plots, with minimal overlap around the origin of the graph. The placement of SIM and Makanjila samples is contained within the range set by the two putative parental populations on axis 1, clustering principally in the area where *M. thapsinogen* and *M. zebra* samples overlap. However, the 2nd FCA axis strongly separates samples of the Makanjila population from the others (axes 3 and 4 contribute no additional separation). The intermediacy of Makanjila samples relative to the putative parental populations, and the similarity of Makanjila samples to the SIM samples is supportive of the hypothesis of hybrid origin. The separation provided by axis 2 of the FCA plots, like the assignment tests, is driven largely by the contribution of the unique Makanjila alleles. Thus, deviations from the pattern of Makanjila intermediacy are generated only by the presence of unique genetic material in the Makanjila population, and not genome-wide differences.

The origin of the unique alleles in the Makanjila population remains unknown. However, they provide evidence that geneflow between the putative parental and hybrid populations has been drastically reduced since the introgression event. That is, even extremely low levels of ongoing geneflow would distribute such high frequency alleles to neighboring populations. Our simulation suggests that their absence in these populations is real.

Discussion

Genetic analysis of the phenotypically intermediate Makanjila *Metriclima* sample shows that this population has arisen as a result of extensive hybridization. The geographical juxtaposition of the Makanjila population suggests that the RT *M. thapsinogen* from Eccles Reef and the BB *M.zebra* from Chiofu Bay may be the parental populations of the extant Makanjila hybrid population. Our microsatellite analyses provide some support for this hypothesis, but indicate that the *Metriaclima* population at Makanjila may have a more complex history.

Private alleles at high frequency and exact tests provide compelling support for the current genetic distinctness of the Makanjila population. In our underwater observations, no pure RT or BB individuals were seen among the fishes at Makanjila, and no intermediate forms were observed at the parental locations. Despite the proximity to the hybrid population, the RT *M. thapsinogen* and BB *M. zebra* populations remain entirely distinct from each other, suggesting little or no backcrossing. Linkage disequilibrium, which would be expected if introgression from the parental populations was ongoing, was not detected in the Makanjila population (data not shown). Together, these data suggest that the initial creation of the hybrid population was followed by a cessation of geneflow, and subsequent isolation of the Makanjila population. This reduction in geneflow might largely have been due to changing water levels that could have physically isolated the population at Makangila. Such a scenario has been proposed as a divergence model for cichlids, wherein lake level dynamics cause diverging populations to become physically isolated accelerating the divergence process (reviewed in Kornfield and Smith 2000).

The presence of unique alleles at Makanjila also bears on our inference of the origin of this hybrid population. The origin by mutation of new allelic variants in hybrid populations, i.e. hybrizymes, is a recognized, though poorly understood phenomenon (Hoffman & Brown 1995, Schilthuizen et al. 2001). Sequence analysis of one allele class (Pzeb4, 126) suggests that it is unlikely that alleles arose in the Makanjila population by mutation since the hybridization event. However, these alleles cannot be explained simply by hybrid origin of the Makanjila population from the flanking populations examined without additional contribution from an unknown population. Below we propose two scenarios for the creation of the Makanjila hybrid population considering the mode of colonization, and incorporation of private alleles. Our data do not discriminate between these two possibilities and the different interpretations do not diminish the importance of our main conclusion: the Makanjila population originated by extensive hybridization and is a genetically independent, distinct hybrid taxon (Allendorf 2001).

The first explanation is constructed wherein the *M. thapsinogen* from Eccles Reef and *M. zebra* from Chiofu Bay are not the direct parents of the Makanjila hybrids. Instead, a distinct RT or BB population closely related to its like-colored neighbor would have existed at Makanjila prior to introgression from the other adjacent population. In this scenario, only one of the two pure populations examined here would have been involved in hybridization. The similarity of the second population to the extant Makanjila population would be a relic of their historical relationship. Unique alleles detected in the hybrids would have been present in the original, pure (RT or BB) Makanjila resident population.

The second scenario that we propose involves colonization of Makanjila by hybrids produced between the *M. thapsinogen* (Eccles Reef) and *M. zebra* (Chiofu Bay) populations. The shallow water at Makanjila suggests that the habitat there is ephemeral, and possibly only recently available. Once submerged due to lake level rise (or sufficiently emerged due to lake level drop), the rocks in this area would become available for habitation, and the first inhabitants would be straying individuals from adjacent populations (McKaye & Gray 1984). The reproductive environment presented to these transient fishes is quite different from that encountered by a migrant into an established population. Indeed, each disperser is presented with a unique 'limited choice' environment, wherein the most suitable mate might be of a different color morph or species. Such a situation is simulated in aquaria, where hybridization is common, even between individuals of different genera (Kornfield & Smith 2000). This breakdown of mate choice behavior due to the environmental conditions explains the lack of geneflow between parental populations.

Under the second scenario, the origin of unique alleles can be explained by introgression from an additional, unidentified population into the Makanjila population after the initial hybrid origin. All four alleles significantly unique to the Makanjila population (with respect to the flanking populations) are found in *Metriaclima* populations throughout the lake at varying frequencies (Van Oppen et al. 2000, personal observations). FCA analysis implicates unique alleles in the separation of the Makanjila population (see Results). Exclusion of just the individuals with a 'Pzeb4-126' allele results in placement of all remaining Makanjila individuals within the cluster formed on axis 1 of the FCA plot, making them indistinguishable from the SIM sample (data not shown). The common alleles at Pzeb3, Pzeb4, and Pzeb5 show a striking pattern of intermediacy with the exception of the unique alleles. Thus, the third population might be a congener with a similar genetic background, but which possesses some genetic divergence. Indeed, two other M*etriaclima* populations exist in the region of Makanjila and they are potential contributors. These populations could not be examined during our fieldwork and have not yet been sampled for genetic analysis. Examination of these populations would be an interesting undertaking to determine whether they might indeed be contributors to the Makanjila hybrid population.

The proposal of multiple hybridization events may not seem parsimonious given the data on low levels of geneflow among adjacent populations (Van Oppen et al. 1997b, Van Oppen et al. 1998, Arnegard et al. 1999, Knight & Turner 1999, Markert et al. 1999, Danley et al. 2000). However, colonization studies on Lake Malawi cichlids have shown rapid habitation of artificial reefs, suggesting extensive movement of many cichlid species (McKaye & Gray 1984). Thus, the low levels of geneflow inferred by population genetics may be caused by biotic factors (such as mate choice) rather than by extrinsic limitations on movement. If the mating tolerances and coloration of the hybrid populations are sufficiently perturbed, geneflow from an additional population might be considerably more likely than one would expect based on the current data.

The discovery of a genetically distinct hybrid population with unique color phenotype is important to cichlid systematists since such populations are frequently elevated to specific status (i.e. Stauffer et al. 1997). The Makanjila population is phenotypically and genetically disntinct, and as such it constitutes a hybrid taxon (Allendorf et al. 2001). As with many of the currently recognized species within the *M.*

zebra complex, the Makanjila population is currently allopatric from congeners and the degree of reproductive isolation cannot be easily evaluated. But to the degree that allopatric populations of *Metriaclima* color variants are considered valid species the Makanjila hybrid population conforms to this standard. While elevation to specific status would certainly be premature,'in this dynamic system this population can be considered an insipient species. Given sufficient time in isolation, this population could evolve on its own trajectory, and might no longer be identifiable as having a hybrid phenotype. The Makanjila population, therefore, appears to be in a transient stage of speciation by hybridization.

Many of the hundreds of currently described species in Lake Malawi are narrow endemics, known only from a single island, reef, or rocky shore within the Lake. Phylogenetic affinities of these endemic species are entirely unknown, and their evolution remains poorly understood. A limited survey of endemics within Lake Malawi showed little or no reduction in genetic diversity when compared to cosmopolitan species in the Lake (Moran & Kornfield 1995), despite narrow distributions and necessarily limited population sizes. Some narrow endemics have color phenotypes intermediate to, or mosaics of, more widely distributed species (e.g. *M. benetos*, *M. greshakei*, *M. mbenjii* (Stauffer et al. 1997)). Taken together, this information suggests that some narrow endemics in Lake Malawi could possibly be derived by hybridization rather than by standard divergence in allopatry or sympatric speciation. This is an extremely compelling possibility despite the uncertainty inherent in such speculation. Identification of species that have potentially arisen via hybridization certainly leads to several testable hypotheses. Increased understanding of the genetic control over various characters of coloration will also aid in our understanding of the potential of hybridization to create new species.

This study provides a new perspective on the maintenance of high levels of genetic variation within cichlid populations. Fluctuations in lake level are thought to continually isolate and merge areas of habitat, facilitating episodic sympatry between divergent allopatric populations, incipient species, or species (Sturmbauer 1998, Kornfield & Smith 2000, Sturmbauer et al. 2001). These episodes of sympatry would test the reproductive isolation of previously allopatric populations; if populations maintain reproductive barriers, species are preserved or established. However, some gene exchange might occur in the process, as complete behavioral isolation between divergent cichlid taxa may not evolve in allopatry (Ruber et al. 2001). With geneflow on secondary contact, previously isolated populations could recover genetic diversity lost during periods of finite population size. Hybridization could thus explain the maintenance of excess genetic variation in widely distributed, as well as narrow endemic Lake Malawi cichlids.

The hybrid *Metriaclima* population at Makanjila was initially recognized by its intermediate color pattern. In the absence of such noticeable features, hybrids probably go undetected. We thus suggest that hybridization may be far more pervasive than currently appreciated in this and other cichlid systems, and that impact on genetic variation may be widespread. If incorporated future studies, hybridization may provide explanatory power for current problems in cichlid evolutionary biology, including the origin of endemic species and the unexplained maintenance of genetic diversity.

Chapter 4

EXTREME COLOR CONVERGENCE WITHIN THE *Pseudotropheus (Maylandia) zebra* (= *Metriaclima zebra*) SPECIES COMPLEX OF LAKE MALAWI CICHLIDS IS REVEALED BY A MICROSATELLITE PHYLOGENY

Abstract

We have constructed a microsatellite phylogeny for one group the extremely diverse, rapidly evolving cichlids of Lake Malawi. The evolution of two common color characters was examined among seventeen populations of ten different species in the *Pseudotropheus (Maylandia) zebra* species complex. Convergence of color pattern appears ubiquitous; we show that lateral barring has been lost four times, and a red dorsal fin has been gained at least two times among these taxa. These two color characteristics have evolved independently of each other, resulting in a mosaic of phenotypes. We present evidence for the paraphyly of the widespread 'blue-black' (BB) *P. zebra* and polyphyly of other color forms examined. Further, we found a slight, but significant correlation between geographic and genetic distances among samples. Two possible explanations are advanced for this result (1) interspecific geneflow created close genetic relationships between adjacent populations, or (2) new taxa arose recently on the periphery of the *P. zebra* distribution and thus retain the genetic distance associated with sister taxa. These two possibilities need not be mutually exclusive. Given that geneflow, if present, has been insufficient to homogenize color differentiation among taxa, we suggest that these data reveal a general pattern of speciation in peripheral isolation. This preliminary microsatellite phylogeny provides a framework for the formulation of

hypotheses to examine evolutionary processes among these fishes and demonstrates the utility and limitations of this class of molecular markers for phylogenetic reconstructions.

Introduction

The cichlid fishes of the African great lakes are a model system for the study of evolution (Kornfield & Smith 2000). A particular focus of recent evolutionary studies of this fauna has been examining the role of color in maintaining (Knight & Turner 1999, Knight et al. 1998, McElroy & Kornfield 1991, Seehausen et al. 1997, Seehausen et al. 1998) and generating (Seehausen & van Alphen 1999, Seehausen et al. 1999, Turner 1995) species diversity via sexual selection. Similar efforts have been directed at deciphering the effect of geological features on geneflow and population structure (Arnegard et al. 1999, Danley et al. 2000, Markert et al. 1999, Markert et al. 2001, Van Oppen et al. 1997b) with an especial focus on the effect of lake level fluctuations (Owen et al. 1990, Rossiter 1995, Sturmbauer et al. 2001). In sum, these studies have contributed a great deal to our general understanding of the origin of species in the East African great Lakes of Victoria, Malawi and Tanganyika. However, an enduring obstacle to our comprehension of evolutionary processes among these fishes is the lack of species level phylogenies. Knowledge of such phylogenetic relationships could provide a framework for critical studies addressing taxonomy, biogeography and speciation.

Previous attempts at molecular phylogenetic reconstruction have been impeded by the maintenance of ancestral polymorphism in extant, divergent lineages (Moran & Kornfield 1993) (Reinthal & Meyer 1997)due to incomplete lineage sorting (Parker & Kornfield 1997, Takahashi et al. 2001). Higher level molecular phylogenies have been informative (Meyer 1993, Meyer et al. 1990, Meyer et al. 1994), and provide some of the most important insights into cichlid evolution to date (Kocher et al. 1993). Recently, Amplified Fragment Length Polymorphism (AFLP) data has been employed to successfully resolve relationships among genera of the Lake Malawi mbuna (rock fish) (Albertson et al. 1999); this method shows considerable promise (Markert et al. 2001). In the present study, we use microsatellite DNA markers to examine genetic relationships among species and geographical color variants within a species complex of the mbuna of Lake Malawi.

Pseudotropheus (Maylandia) zebra is an extensively studied Lake Malawi mbuna with a lake wide distribution and extensive color variation among allopatric populations. In a recent taxonomic revision of this group, ten new species were described, some of which differ in morphometric and meristic characters while others differ principally in one or more characters of coloration (Stauffer et al. 1997). Additional color variants exist within this species complex that are yet to be taxonomically considered (Konings 2001). We examine the evolution of two characters of coloration: lateral barring and dorsal fin color. Lateral barring is present or absent, while dorsal fin color among these taxa is either blue, red or black. Among the taxa studied, these two characters exist in nearly every combination resulting in five distinct color phenotypes: BB (common name: 'blue-black'; phenotype: barred flank, blue dorsal fin), RT ('red top'; barred, red dorsal), BD ('black dorsal'; barred, black dorsal), CB ('cobalt'; not barred, blue dorsal) and RTC ('red top cobalt'; not barred, red dorsal). Multiple allopatric populations of each phenotype are represented here (Fig 1).

The general objective of this species level phylogenetic reconstruction is to provide a framework for evolutionary studies on the diverse and well studied *Pseudotropheus (Maylandia) zebra* species complex in order to evaluate biogeographic and speciation hypotheses. Using this phylogeny we evaluate color phenotypes. Our findings could potentially reveal trends in the evolution of coloration in cichlids that might be applicable to other genera in Lake Malawi as well as in other African Lakes. Indeed, factors of central importance such as trophic adaptations (Kocher et al. 1993), habitat fidelity (Markert et al. 1999, Sturmbauer et al. 1997), parental care (Goodwin et al. 1998), sexual selection (Seehausen & van Alphen 1999, Turner 1995) and the effect of lake level fluctuations (Sturmbauer 1998, Sturmbauer et al. 2001) have all unified the phylogenentically independent radiations of the African Great Lakes. Evolution of color pattern within a species complex might be a unifying element in the evolution of African cichlids.

Materials and methods

Study species

The *Pseudotropheus (Maylandia) zebra* species complex has been the subject of taxonomic uncertainty. Within *Pseudotropheus* the subgenus *Maylandia* was created for this group (Meyer & Foerster 1984). However, the genus *Metriaclima* was later erected for this species complex, excluding the name *Maylandia* as the generic designator (Stauffer et al. 1997). In the current taxonomy, *Metriaclima* is considered a junior synonym of *Maylandia* (Conde & Gery 1999). We follow other recent studies (Rico &

Turner 2002) in using the former, valid *Psudotropheus* (*Maylandia*) generic assignment until this taxonomic issue is fully resolved.

Seventeen populations belonging to ten formally described species and one undescribed taxon within the *Pseudotropheus (Maylandia) zebra* species complex were sampled from rocky habitat on the shores and islands of Lake Malawi (Fig. 4.1, Table 4.1). Each sampled population is characterized as having one of the five (two character) phenotypes (described above), and each phenotype was sampled in a minimum of two locations. Several of our samples are narrow endemics, and were thus sampled from their entire known range (Stauffer et al. 1997). Some samples were previously examined in a phylogeographic study of RT and BB forms (Smith & Kornfield, manuscript in review). Location and phenotypic identity of samples is shown in Fig 4.1.

Sampling & DNA preparation

Fish were collected with monofilament gill nets using SCUBA. Only males were sampled to ensure proper species identification. Small portions of the caudal fin were clipped from individual fish and preserved in 95% ethanol; ethanol was changed after 48 hours. DNA was extracted using the QiAmp tissue kit (Qiagen, Valencia, California) following the manufacturer's instructions.

Table 4.1. Location and genetic diversity of population samples.

			Location					
Letter	Taxon	Name	S Longitude	E Latitude	N	He	Ho	A
a	*P. zebra*	Nkhata Bay	11:36.370	34:17.992	71	0.8143	0.8284	21.50
b	*P. callainos*	Nkhata Bay	11:36.370	34:17.992	71	0.8270	0.8326	19.63
c	*P.* sp. 'zebra chilumba'	Luwino	10:26.775	34:16.996	55	0.8335	0.7673	23.50
d	*P.* sp. 'zebra chilumba'	Mpanga Rocks	10:25.779	34:16.720	37	0.8139	0.7424	18.75
e	*P. emmiltos*	Mpanga Rocks	10:25.779	34:16.720	70	0.8239	0.8025	20.88
f	*P. zebra*	Taiwan Reef	11:57.334	34:35.288	42	0.7848	0.7844	17.50
g	*P. zebra*	Chizumulu Island	12:01.367	34:37.745	42	0.8178	0.8214	20.25
h	*P. zebra*	Likoma Island	12:06.576	34:43.599	73	0.8020	0.7620	20.75
i	*P. estherae*	Minos Reef	12:53.213	34:45.013	48	0.7780	0.7388	20.13
j	*P. zebra*	Chiofu Bay	13:31.773	34:51.962	70	0.8279	0.8273	21.50
k	*P. thapsinogen*	Eccles Reef	13:45.886	34:56.251	72	0.8155	0.8083	20.00
l	*P. benetos*	Mazinzi Bay	14:07.846	34:56.644	70	0.8478	0.8714	20.00
m	*P. sandaracinos*	Namaso Bay	14:09.680	35:00.000	67	0.7713	0.7687	20.63
n	*P. grashekaei*	Makakola Reef	14:15.482	35:07.545	50	0.8588	0.8499	18.25
o	*P. pyrsonotos*	Nakantenga Reef	13:55.218	34:38.546	71	0.8519	0.8468	22.13
p	*P. zebra*	Namalenji Island	13:43.754	34:38.462	64	0.7814	0.7823	16.75
q	*P. mbenjii*	Mbenji Island	13:25.760	34:28.141	35	0.8102	0.8321	19.38

Number of samples (N), Expected Heterozygosity (He), Observed Heterozygosity (Ho), Mean Number of allleles across all loci (A).

Eight microsatellite loci were examined: Pzeb2, Pzeb3, Pzeb4, Pzeb5, (Genbank accession numbers X99783, X99784, X99786, X99787 (Van Oppen et al. 1997a)) UME002, UME003, (U14396, U14397 (Parker & Kornfield 1996)), UNH231 (G12382 (Kocher et al., unpublished)) and DXTUCA3 (U94850 (Seultmann et al., unpublished)). For each sample, alleles at the eight loci were amplified by the polymerase chain reaction (PCR) using standard conditions and separated on an ABI 377 DNA sequencer as per the

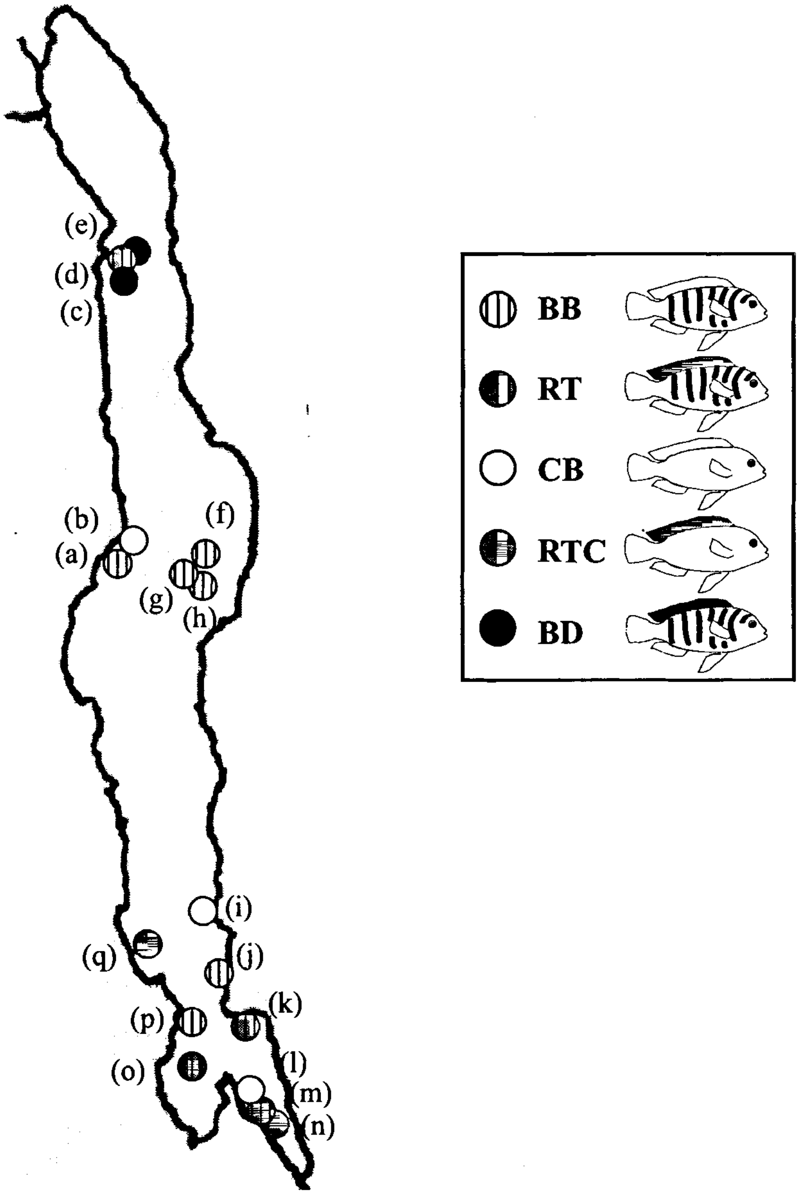

Figure 4.1. Sampling locations for seventeen populations with five distinct color phenotypes from ten species of the *Pseudotropheus (Maylandia) zebra* species complex. Each sample is represented by a letter (lowercase, in parentheses). Sample markers are patterned to represent each two character color phenotype, as detailed in the key. Diagrams of each color phenotype are also represented in the key, and subsequently referenced in Fig. 4.2. See text for descriptions of each color phenotype. Taxonomic designation, sampling location, and genetic diversity of each sample are displayed in Table 4.1.

manufacturer's instructions (Applied Biosystems International). PCR fragments were detected using Genescan (v 3.1) software, and alleles were scored using Genotyper (v 2.1) software (Applied Biosystems International).

Data analysis

Genetic distance between populations was measured with Nei's standard genetic distance (Ds) (Nei 1972). In an empirical study of genetic distances, Paetkau et al. (1997) concluded that Ds was an effective metric for reconstructing relationships between populations that diverged in the last 20 kyrs. Further, Nei's standard distance is based on the infinite alleles model of microsatellite mutation (IAM), and does not require that microsatellites conform to a strict stepwise mutation model. Genetic distance was calculated using the GENDIST module of PHYLIP, and the neighbor-joining (extended majority rule, consensus) dendogram was generated using the NEIGHBOR and CONSENS modules of PHYLIP (Felsenstein 1993). Data were re-sampled over loci to generate bootstrap support.

Because deep open water and sandy shoreline limits mbuna migration (Markert *et al.* 1999) populations separated by very short distances of these substrates may be more likely to exchange migrants than those separated by longer distances of mbuna habitat. However, because of the large geographical range between samples in this study and because of the existence of a paleo-shoreline that has moved in the past (Johnson 1996, Owen et al. 1990, Sturmbauer et al. 2001) we have chosen to measure geographic distance using the spherical distance implemented in the Great Circle Calculator (www.waypoint.org). While it may not accurately represent the possibility of migration between any two taxa, it simply characterizes an assumption-free measure of gross distance between samples. Correlation between geographic and genetic distance was measured using a Mantel test.

Results

Color convergence within this species complex is ubiquitous. We constructed a neighbor joining consensus tree for the seventeen studied populations using the extended majority rule (Fig. 4.2). Nodes supported by greater than 50% of the bootstrap replicates are indicated; bootstrap support is moderate, but uniform.

Because few microsatellite phylogenies have been attempted (Petren et al. 1999), no reliable context exists for quantifying the strength of a tree. Given bootstrap resampling over the limited set of loci used (n=8), an uninformative locus might substantially decrease average bootstrap values, and accepted bootstrap thresholds may not be appropriate. To gauge the significance of our bootstrap values we compared them to those associated with randomly generated data. Individuals were globally permuted to create seventeen random population samples, and used to construct fully resolved (14 nodes) consensus trees from which mean bootstrap values were calculated. We generated a distribution of mean bootstrap values for ten such trees. The mean bootstrap value for the fully resolved consensus tree from the actual data (mean = 44.10) was significantly different (t=4.59, df=9, p=0.001) from the distribution of values from the randomly generated data (mean=26.75, s.d.= 3.6).

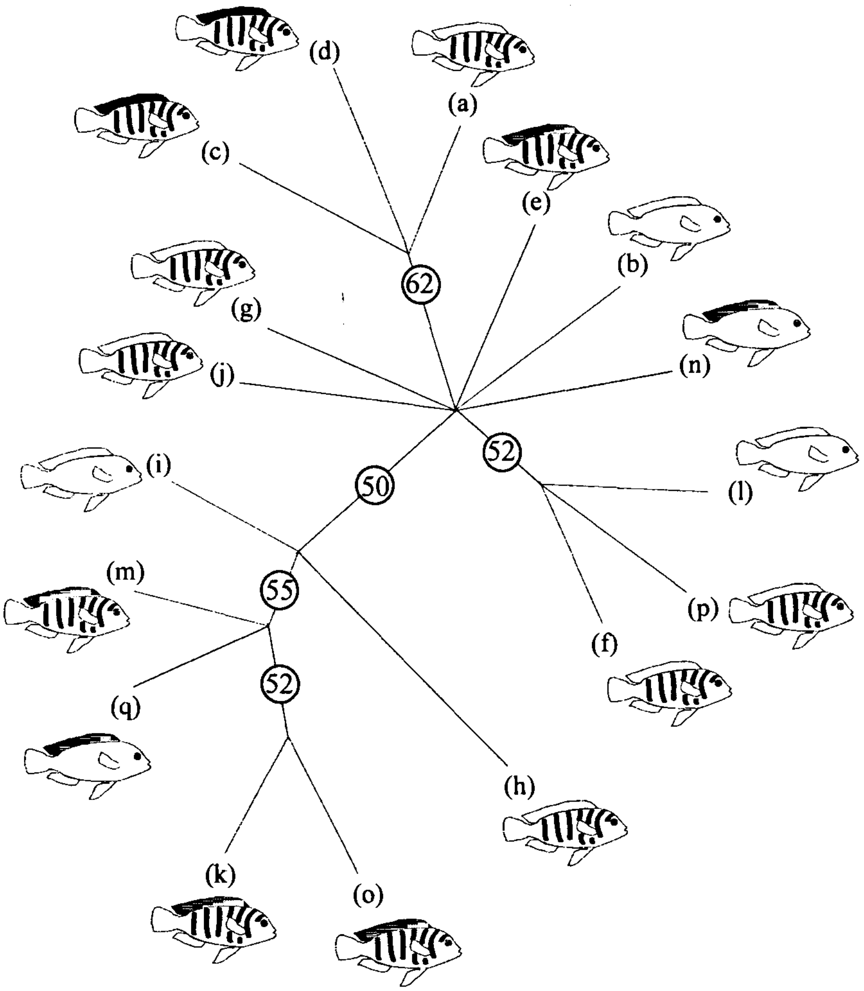

Figure 4.2. An unrooted phylogeny for the seventeen taxa in the *Pseudotropheus (Maylandia) zebra* species complex showing extensive convergence in color phenotype. Bootstrap support for nodes supported by greater than 50% of the resampled data sets is displayed in a circle on each node. Diagrams of each color phenotype as shown in Fig. 4.1 are used to represent taxa and letters correspond to samples as detailed in Fig. 4.1 and Table 4.1. The BB taxa (a, f, g, h, j, p) are paraphyletic, RT taxa (e, k, m, o) are polyphyletic, CB taxa (i, l) have two distinct origins, RTC taxa (n, q) have two distinct origins, and the two BD taxa from Chilumba (c, d) cluster together. The two color characters have evolved independently of each other. In the most parsimonious scenario, bars have been lost four times (i, q, l and b/n) and the red dorsal fin has been gained twice (k/m/o/q and e/n), while the black dorsal fin has a single origin.

Owing to limitations on sampling isolated populations and microsatellite genotyping (>8000 individual genotypes were examined in this study) some of these loci may not have been sufficiently characterized, and may distort the signal to noise ratio, thus lowering bootstrap values. As a preliminary investigation of the influence of allelic diversity on phylogenetic signal, we devided the data into high (>40 alleles; PZEB2, UNH231, UME002, UME003) and low (< 30 alleles; PZEB3, PZEB4, PZEB5, DXTUCA3) diversity sets and analyzed them separately. The low diversity loci produced the same tree topology as the full dataset and mean bootstrap support was significantly different from the distribution of random data (mean=40.4, p= 0.01). The high diversity loci produced a markedly different tree topology and mean bootstrap support that was not significantly different from the random data. This result hints at substantial noise in these data and indicates that some loci might be more reliable than others for phylogenetic reconstructions.

Because of the lack of previous phylogenetic knowledge for this group and the problems associated with resolving distant relationships using microsatellites, the phylogenetic tree is shown unrooted. The CB (not barred, blue dorsal) and RTC (not barred, red dorsal) forms are here found to be polyphyletic. As noted in a previous examination of these taxa, the BB (barred, blue dorsal) form is paraphyletic and the RT (barred, red dorsal) is polyphyletic (Smith and Kornfield, manuscript in review). Each color character state (except black dorsal fin, which had a limited geographic sample) has originated independently multiple times. Lateral barring has been lost four times among these seventeen taxa and the red dorsal fin has at least two distinct origins.

A slight (Mantel test; $r^2 = 0.1151$, $g = 3.00$), but highly significant ($p < 0.01$) correlation was detected between genetic distance and geographic distance among all taxa. This correlation is typical of conspecific populations that exist along a chain of unbroken geneflow (e.g. Paetkau et al. 1997), but is highly unexpected here, in an examination of populations from ten distinct species. Among conspecific BB *P. zebra* samples, a distance correlation was also detected ($r^2 = 0.192$, $g = 1.902$, $p < 0.05$).

Discussion

This microsatellite phylogeny indicates that color convergence within *Pseudotropheus (Maylandia)* is widespread. One immediate implication of this result is that the nearly inestimable color diversity of cichlid fishes might actually under represent true species diversity.

The two color characters examined here appear to evolve independently of each other resulting in a mosaic of character states in extant taxa. In a quantitative genetic study, Danley (2001) found that lateral barring and fin coloring in some *Pseudotropheus (Maylandia)* species are controlled by a few loci and are genetically independent (Danley 2001). Thus, combinations of a small number of color characters having even a few character states can potentially produce a very large number of composite color phenotypes. Many other genera of mbuna as well as taxa in Lake Victoria also express barring and dorsal fin variation among their species (Konings 2001). Our results are thus generally applicable to cichlid radiations in African lakes.

The observed correlation between geographic and genetic distance across all samples is highly unexpected. In fact, spherical distances calculated for species pairs on opposite sides of the Lake separated by extremely deep waters (ie. Fig. 4.1, a,b/g,h,f) should substantially reduce such associations. At least two explanations can be advanced to explain this correlation: (1) geneflow across species boundaries, and (2) phylogenetic signal associated with geographical origin. Hybridization has been observed in all three East African Great Lakes (Ruber et al. 2001, Salzburger et al. 2002, Seehausen et al. 1997, Stauffer et al. 1996), and multiple times within *Pseudotropheus (Maylandia)* (Smith and Kornfield, manuscript in review, Konings, personal communication, Stauffer & Hert 1992). Hybridization may be an ongoing process, or it may occur episodically during periodic lake level fluctuations. However, it is unlikely that hybridization would be sufficient to erase phylogenetic signal at microsatellite markers while taxa remained phenotypically distinct. Thus, phylogenetic signal is likely maintained, even though hybridization may contribute to the geographical association among genetic distances.

The distance correlation, together with this microsatellite phylogeny reveals a model for the evolution of the allopatric color diversity in this species complex. The maintenance of the correlation between geographic and genetic distance among taxa through speciation events underscores the possibility of a tremendously recent origin for this sub-generic diversity. The most widespread taxon in this study is *P. zebra* (BB), which has a nearly lake wide distribution and exhibits isolation-by-distance. The patchy distributions of the other sampled taxa generally interrupt the *P. zebra* distribution on short stretches of rocky coastline or islands. Narrow distributions of color differentiated taxa within the *P. zebra* range have been interpreted as consistent with microallopatric speciation (Fryer 1959). Paraphyly of BB *P. zebra* populations observed here is consistent with these patchily distributed taxa having arisen recently from the *P. zebra* distribution in peripheral isolation, possibly via microallopatric speciation or "divergence with geneflow" (Danley et al. 2000). Concordant with previous conclusions (Smith & Kornfield, in review), this interpretation suggests that some dispersal has occurred since divergence.

Clustering of the southern RT taxa could be explained by dispersal after divergence from the widespread BB form. Taken together with our phylogenetic data, species distributions suggest a general trend of origination of each taxon in peripheral isolation followed by range expansion. The biological process that causes taxa to diverge in color remains unknown. Our phylogeny suggests that divergence from the widely distributed BB *P. zebra* might be followed by dispersal and subsequently, further divergence of color. If this repetitive process operates via natural or sexual selection, genetic variation underlying coloration traits might be exhausted. However, because there may be several genetically independent quantitative traits under selection (Danley 2001), additive genetic variation could be maintained during recurrent rounds of selection (Brooks & Endler 2001).

Our results, like those of Albertson et al. provide no evidence for sympatric speciation in the southeast arm of the Lake. We sampled three distinct taxa from this region (CB(l), RT(m) and RTC(n)), all of which appear to have different ancestry. Albertson et al. examined sympatric *P. zebra* and *P. benetos* (RTC(n)) populations also from the southeast arm of Lake Malawi and determined that they were not sister taxa, and thus did not diverge in sympatry (Albertson et al. 1999). Our results are concordant with theirs and further indicate that *P. benetos* has not diverged from the other *Pseudotropheus* (*Maylandia*) species sampled from the southeast. Instead it shares ancestry with a more distant population of *P. zebra*. Because the habitat in this region of the lake is only newly available to mbuna due to lake level rise (Johnson 1996, Owen et al. 1990), the taxa present must have either evolved recently in situ or colonized from deeper portions of the lake. Our data provides evidence that the diversity of this region may be best described by "progressive invasion" of divergent lineages (Bouton 2000), though additional color divergence in the south may have occurred.

Microsatellite DNA markers have been used extensively in studying the differentiation of and relationships among mbuna populations (Arnegard et al. 1999, Danley 2001, Danley et al. 2000, Knight et al. 1998, Knight et al. 1999, Kornfield & Parker 1997, Markert et al. 1999, Markert et al. 2001, Rico & Turner 2002, Van Oppen et al. 1997a, Van Oppen et al. 1997b). Comparison across species boundaries has resulted here in successful phylogenetic reconstruction. The examination of trees generated from random data (see Results) suggests that phylogenetic signal does indeed exist in the microsatellite allele frequency distributions among these taxa. The reconstructed phylogeny is concordant with geography and is biogeographically plausable (discussed above), strongly indicating non-random tree structure. The taxonomic study of Stauffer et al. (1997) created specific designations using coloration as well as morphological data. The species constructed in that study predict many of the relationships represented here, indicating morphological concordance with our microsatellite phylogeny. Finally, our phylogeny accurately represents well studied relationships among taxa. For example, the sympatric species pair at Nkhata Bay (*P. zebra* and *P. callainos*) are known to be reproductively isolated (Knight & Turner 1999, Knight et al. 1998) and phylogenetically distinct, our tree is in agreement with this data. Similarly, the *P. sp.* 'zebra chilumba' populations from Mphanga Rocks and Luwino are considered conspecific and are endemic to that restricted region (Konings 2001), and our phylogeny also accurately reconstructs this relationship.

Conclusions

The microsatellite phylogeny erected here contributes to studies demonstrating that these molecular markers are useful phylogenetic tools for some recently evolved taxa (Harr et al. 1998, Petren et al. 1999). The phylogenetic patterns resolved indicate an extreme level of color convergence among congeneric taxa. Biogeographical interpretation of this phylogeny suggests that the *Pseudotropheus (Maylandia)* species that interrupt the *P. zebra* distribution may have diverged within the *P. zebra* range, possibly by microallopatric speciation. Further, the phylogeny suggests that dispersal and additional species divergence may have occurred subsequent to the radiation of lineages from *P. zebra*. The phylogenetic hypothesis here erected for the *Pseudotropheus (Maylandia) zebra* species complex provides a framework for experimental examination of the processes that might contribute to the origin of species among African cichlids.

LITERATURE CITED

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

Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001. The problems with hybrids: setting conservation guidelines. *Trends Ecol. Evol.* 18: 613-622

Andersson M. 1994. *Sexual Selection*. Princeton: Princeton University Press. 599

Arnegard ME, Markert JA, Danley PD, Stauffer JR, Ambali AJ, et al. 1999. Population structure and colour variation of the cichlid fish *Labeotropheus fuellerborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proc. R. Soc. Lond. B* 266:119-130

Arnold ML. 1997. *Natural Hybridization and Evolution*. New York: Oxford Press. 256

Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JJ. 1999. Natural hybridization: how long can you go and still be important? *Ecology* 80:371-381

Avise JC. 1994. *Molecular Markers, Natural History and Evolution*. New York: Chapman and Hall. 511

Banks M, Eichert W.1999. WhichRun version 4.1. Available from http://www-bml.ucdavis.edu/whichrun.htm.

Barraclough T, Harvey P, Nee S. 1995. Sexual selection and taxonomic diversity of passerine birds. *Proc. R. Soc. Lond. B* 259:211-215

Barraclough T, Vogler AP, Harvey PH. 1998. Revealing the factors that promote speciation. *Phil. Trans. R. Soc. Lond. B* 353:241-249

Belkhir, K. 1999. GENETIX V. 4.0. Montpellier, France: Laboratoire Genome, Populations, Interactions, CNRS UPR 9060

Bouton, N. 2000. Progressive invasion and allopatric speciation can also explain distribution patterns of rock-dwelling cichlids from southern Lake Victoria: a comment on Seehausen and van Alphen (1999). *Ecol. Lett.* 3:166-171

Bowers N, Stauffer JR, Kocher TD. 1994. Intra- and interspecific mitochondrial DNA sequence variation within two species of rock-dwelling cichlids. *Mol. Phyl. Evol.* 3:75-82

Brooks R, Endler J. 2001. Direct and indirect sexual selection and quantitative genetics of male traits in guppies (*Poecilia reticulata*). *Evolution* 55:1002-1015

Cannatella DC, Hillis DM, Chippindale PT, Weigt L, Rand AS, et al. 1998. Phylogeny of frogs of the *Physalaemus Pustulosus* species group, with an examination of data incongruence. *Syst. Biol.* 47:311-335

Carleton KL, Hárosi FI, Kocher TD. 2000. Visual pigments of African cichlid fishes: evidence for ultraviolet vision from microspectrophotometry and DNA sequences. *Vision Res.* 40:879-890

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

Clark AG. 1997. Neutral behavior of shared polymorphism. *Proc. Natl. Acad. Sci. USA* 94:7730-7734

Conde B, Gery J. 1999. *Maylandia* Meyer et Foerster, 1984, un nom generique disponible (Teleostei, Perciformes, Cichlidae). *Revue Francaise d' Aquariologie et de Herpetologie* 26:21-22

Coyne J, Orr HA. 1998. The evolutionary genetics of speciation. *Phil. Trans. R. Soc. Lond. B* 353:287-305

Crapon de Caprona MD, Fritzsch B. 1984. Interspecific fertile hybrids of haplochromine Cichlidae (Teleostei) and their possible importance for speciation. *Neth. J. Zool.* 34:503-538

Danley P. 2001. The ecological, behavioral, and genetic factors influencing the diversification of Lake Malawi's rock-dwelling cichlids (East Africa). Durham, New Hampshire: University of New Hampshire

Danley PD, Kocher TD. 2001. Speciation in rapidly diverging systems: lessons from Lake Malawi. *Molecular Ecology* 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

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

Dominey W. 1984. Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian *Drosophila*. In: Echelle AA, Kornfield I, eds. *Evolution of Fish Species Flocks*. Orono, Maine: University of Maine Press. 231-249.

Dowling TE, Secor CL. 1997. The role of hybridization and introgression in the diversification of animals. *Annu. Rev. Ecol. Syst.* 28:593-619

Eccles DE, Trewavas E. 1989. *Malawian Cichlid Fishes: the Classification of Some Haplochromine Genera*. Herten: Lake Fish Movies. 334

Farias IP, Orti G, Sampaio I, Schneider H, Meyer A. 1998. Mitochondrial DNA phylogeny of the family Cichlidae: monophyly and fast molecular evolution of the neotropical assemblage. *J. Mol. Evol.* 48:703-711

Felsenstein J. 1993 PHYLIP (phylogeny inference package) v. 3.5c. Available from: http://evolution.geentics.washington.edu/phylip.

Finney BP, Scholz CA, Johnson TC, Trumbore S, Southon J. 1996. Late Quaternary lake-level changes of Lake Malawi. In: TC Johnson, EO Odada, eds. *The Limnology, Climatology, and Paleoclimatology of the East African Lakes*. Toronto: Gordon and Breach. p. 495-508.

Foster SA, Scott R, Cresko WA. 1998. Nested biological variation and speciation. *Phil. Trans. R. Soc. Lond. B* 353:207-218

Fournieret Y, Plisnier PD, Micha JC. 1992. Feeding habits of four species belonging to the genus *Haplochromis* (Teleostei, Cichlidae) in Ihema Lake (Rwanda). *Ann. Limnologie* 28:58-69

Fryer G. 1959. The trophic interrelationships and ecology of some littoral communities of Lake Nyasa with especial reference to the fishes, and a discussion of the evolution of a group of rock-frequenting Cichlidae. *Proc. Zool. Soc. Lond.* 132:153-281

Fryer, G. 2001 On the age and origin of the species flock of Haplochromine cichlid fishes of Lake Victoria. *Proc. R. Soc. Lond. B* 268:1147-1152

Fryer G, Iles TD. 1972. *The Cichlid Fishes of the Great Lakes of Africa*. London: Oliver and Boyd. 641

Galis F, Metz J. 1998. Why are there so many cichlid species? *Trends Ecol. Evol.* 13:1-2

Goldschmidt T. 1996. *Darwin's Dreampond.* Cambridge, Massachusetts: MIT Press. 274

Goodwin NB, Balshine-Earn S, Reynolds JD. 1998. Evolutionary transitions in parental care in cichlid fish. *Proc. R. Soc. Lond. B.* 265:2265-2272

Grant PR, Grant BR. 1997. Genetics and the origin of bird species. *Proc. Natl. Acad. Sci. USA* 94:7768-7775

Greenwood PH. 1965a. The cichlid fishes of Lake Nabugabo, Uganda. *Bull. Br. Mus. Nat. Hist. Zoo.l* 12:315-357

Greenwood PH. 1965b. Environmental effects on the pharyngeal mill of a cichlid fish. *Proc. Linn. Soc. Lond.* 176:1-10

Greenwood PH. 1974. The cichlid fishes of Lake Victoria, East Africa: the biology and evolution of a species flock. *Bull. Br. Mus. Nat. Hist. Zool* :1-134

Greenwood PH. 1984. African cichlids and evolutionary theories. In: Echelle AA, Kornfield I, eds. *Evolution of Fish Species Flocks*. Orono, Maine: University of Maine Press. p. 141-154

Greenwood PH. 1991. Speciation. In: Keenlyside M. ed. *Cichlid Fishes: Behaviour, Ecology, and Evolution*. London: Chapman and Hall. p. 84-102

Harr B, Weiss S, David J, Brem G, Schlötterer C. 1998 A microsatellite-based multilocus phylogeny of the *Drosophila melanogaster* species complex. *Curr. Biol.* 8:1183-1186

Harrison R. 1998. Linking evolutionary pattern and process; the relevance of species concepts for the study of speciation. In: Howard DH, Berlocher SH, eds. *Endless Forms Species and Speciation*. New York: Oxford University Press. p. 19-31

Hedrick, P. W. 1999 *Genetics of Populations*. Sudbury, MA: Jones and Bartlett. 553

Hert E. 1990. Factors in habitat partitioning in *Pseudotropheus aurora* (Pisces: Cichlidae), an introduced species to a species-rich community of Lake Malawi. *J. Fish Biol.* 36:853-857

Hert E. 1991. Female choice based on egg-spots in *Pseudotropheus aurora* Burgess 1976, a rock-dwelling cichlid of Lake Malawi, Africa. *J. Fish Biol.* 38:951-953

Higashi M, Takimoto G, Yamamura N. 1999. Sympatric speciation by sexual selection. *Nature* 402:523-526

Hoelzer GA, Melnick DJ. 1994. Patterns of speciation and limits to phylogenetic resolution. *Trends Ecol. Evol.* 9:104-105

Hoffman SMG, Brown WM. 1995. The molecular mechanism underlying the "rare allele phenomenon" in a subspecific hybrid zone of the California field mouse, *Peromyscus californicus*. *J. Mol. Evol.* 41:1165-1169

Holzberg S. 1978. A field and laboratory study of the behaviour and ecology of *Pseudotropheus zebra* (Boulenger), and endemic cichlid of Lake Malawi (Pisces: Cichlidae). *Z. zool. Syst. Evolut.-forsch* 16:171-187

Johnson P, Gulberg U. 1998. Theory and models of sympatric speciation. In: Howard D, Berlocher S, eds. *Endless Forms: Species and Speciation* New York: Oxford University Press. p. 79-89

Johnson TC. 1996. Sedimentary processes and signals of past climatic change in the large lakes of the East African rift valley. In: Johnson TC, Odada EO, eds. *The Limnology, Climatology and Paleoclimatology of the East African Lakes.* Toronto: Gordon and Breach. p. 367-412

Karino K. 1996. Tactic for bower acquisition by male cichlids, *Cyathopharynx furcifer*, in Lake Tanganyika. *Ichthyol. Res.* 43:125-132

Karino K. 1997. Female mate preference for males having long and symmetric fins in the bower-holding cichlid *Cyathopharynx furcifer. Ethology* 103:883-892

Kaufman L. 1997. Asynchronous taxon cycles in haplochromine fishes of the greater Lake Victoria Region. *S. Afr. J. Sci.* 93:601-606

Kaufman LS, Ochumba P. 1993. Evolutionary and conservation biology of cichlid fishes as revealed by faunal remnants in Lake Victoria. *Conserv. Biol.* 7:719-730

Kawanabe H, Hori M, Nagoshi M, eds. 1997. *Fish Communities in Lake Tanganyika.* Sakyo-ku, Japan: Kyoto University Press. 298

Kelly JK, Noor MAF. 1996. Speciation by reinforcement: a model derived from studies of *Drosophila. Genetics* 143:1485-1497

Klein D, Ono H, O'hUigín C, Vincek V, Goldschmidt T, et al. 1993. Extensive MHC variability in cichlid fishes of Lake Malawi. *Nature* 364:330-334

Klein J, Sato A, Nagl S, O'hUigín C. 1998. Molecular trans-species polymorphism. *Ann. Rev. Ecol. Syst.* 29:1-21

Knight ME, Turner GF. 1999. Reproductive isolation among closely related Lake Malawi cichlids: can males recognize conspecific females by visual cues? *Anim. Behav.* 58:761-768

Knight ME, Turner GF, Rico C, VanOppen MJH, Hewitt GM. 1998. Microsatellite paternity analysis on captive Lake Malawi cichlids supports reproductive isolation by direct mate choice. *Mol. Ecol.* 7:1605-1610

Knight ME, VanOppen MJH, Smith HL, Rico C, Hewitt GM, et al. 1999. Evidence for male-biased dispersal in Lake Malawi cichlids from microsatellites. *Mol. Ecol.* 8:1521-1527

Kocher TD, Conroy JA, McKaye KR, Stauffer JR. 1993. Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. *Mol. Phylo. Evol.* 2:158-165

Kocher TD, Conroy JA, McKaye KR, Stauffer JR, Lockwood SF. 1995. Evolution of NADH dehydrogenase subunit 2 in East African cichlid fish. *Mol. Phylo. Evol.* 4:420-432

Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B. 1998. A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* 148:1225-1232

Kondrashov A, Shpak M. 1998. On the origin of species by means of assortative mating. *Proc. R. Soc. Lond. B* 265:2273-2278

Konings, A. 2001 *Malawi Cichlids in Their Natural Habitat.* El Paso, Texas: Cichlid Press. 351

Kornfield I. 1978. Evidence for rapid speciation in African cichlid fishes. *Experientia* 34:335-336

Kornfield I. 1991. Genetics. In: Keenlyside M, ed. *Cichlid Fishes: Behaviour, Ecology, and Evolution*. London: Chapman and Hall. p. 103-128

Kornfield I, Parker A. 1997. Molecular systematics and a rapidly evolving species flock: the *mbuna* of Lake Malawi and the search for phylogenetic signal. In: Kocher TD, Stepien CA, eds. *Molecular Systematics of Fishes*. San Diego: Academic Press. p. 25-37

Kornfield I, Reinthal PN, Lobel PS, Merrill S. 1991. Selection and constraints on bower morphology in a Malawi cichlid. In: *71st Meeting American Society of Ichthyology and Herpetology* AMNH, New York

Kornfield I, Smith PF. 2000 African Cichlid fishes: models systems for evolutionary biology. *Annu. Rev. Ecol. Syst.* 31:163-196

Kornfield I, Taylor JN. 1983. A new species of polymorphic fish, *Cichlasoma minckleyi* from Cuatro Cienegas, Mexico, (Teleostei: Cichlidae). *Proc. Biol. Soc. Wash.* 96:253-269

Kosswig C. 1947. Selective mating as a factor for speciation in cichlid fish of East African lakes. *Nature* 159:604-605

Kruuk L. 1999. Sticklers for sympatry. *Trends Ecol. Evol.* 14:465-466

Kuusipalo L. 1998. Scale morphology in Malawian cichlids. *J. Fish Biol.* 52:771-781

Ladiges W. 1968. Die bedutung okologischer faktoren fur die differenzierung der cichliden Tanganjika- and des Njassa-Sees. *Int. Rev. ges. Hydrobiol.* 53:339-352

Lande R. 1981. Models of speciation by sexual selection on polygenic traits. *Proc. Nat. Acad. Sci. USA* 78:3721-3725

Lande R. 1982. Rapid origin of sexual isolation and character divergence in a cline. *Evolution* 36:213-223

Lazzaro X. 1991. Feeding convergence in South American and African zooplanktivorous cichlids *Geophagus brasilensis* and *Tilapia rendalli*. *Env. Biol. Fishes* 31:283-293

Lessa EP, Cook JA. 1998. The molecular phylogenetics of tuco-tucos (genus *Ctenomys*, Rodentia: Octodontidae) suggests an early burst of speciation. *Mol. Phylo. Evol.* 9:88-99

Lewis DS. 1982. A revision of the genus *Labidochromis* (Teleostei: Cichlidae) from Lake Malawi. *Zool. J. Linn. Soc.* 75:189-265

Lewis PO, Zaykin D. 2001 Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html.

Lippitsch E. 1993. A phyletic study on lacustrine haplochromine fishes (Perciformes, Cichlidae) of East Africa, based on scale and squamation characters. *J. Fish Biol.* 42:903-946

Lippitsch E. 1998. Phylogenic study of cichlid fishes in Lake Tanganyika: a lepidological approach. *J. Fish Biol.* 53:752-766

Luikart G, Sherwin B, Steele B, Allendorf F. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Mol. Ecol.* 7:963-974

Magurran A. 1998. Population differentiation without speciation. *Phil. Trans. R. Soc. Lond. B* 353:275-286

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

Markert JA, Danley PD, Arnegard ME. 2001. New markers for new species: microsatellite loci and the East African cichlids. *Trends Ecol. Evol.*16:100-107

Martens K, Godderis B, Coulter G, eds. 1994. *Speciation in Ancient Lakes.* (Advances in Limnology, Vol. 44). Port Jervis, New York: Lubrecht and Cramer Ltd. 508.

Mayr E. 1963. *Animal Species and Evolution.* Cambridge: Harvard Univ. Press. 811.

Mayr E. 1984. Evolution of fish species flocks: a commentary. In: Echelle AA, Kornfield I, eds. *Evolution of Fish Species Flocks.* Orono, Maine: University of Maine Press. p. 3-12

McCune A, Lovejoy N. 1998. The relative rate of sympatric and allopatric speciation in fishes. In: Howard D, Belocher S, eds. *Endless Forms: Species and Speciation.* New York: Oxford University Press. p. 172-185

McElroy DM, Kornfield IL. 1990. Sexual selection, reproductive behavior, and speciation in the *mbuna* species flock of Lake Malawi (Pisces: Cichlidae). *Env. Biol. Fishes* 28:273-284

McElroy DM, Kornfield IL. 1993. Novel jaw morphology in hybrids between *Pseudotropheus zebra* and *Labeotropheus fullerborni* (Teleostei: Cichlidae) from Lake Malawi, Africa. *Copeia* 1993:933-945

McElroy DM, Kornfield IL, Everett J. 1991. Coloration in African cichlids: diversity and constraints in Lake Malawi endemics. *Nether. J. Zool.* 41:250-268

McKaye KR. 1991. Sexual selection and the evolution of the cichlid fishes of Lake Malawi, Africa. In: Keenlyside M, ed. *Cichlid Fishes: Behaviour, Ecology, and Evolution.* London: Chapman and Hall. p. 241-257

McKaye K, Gray N. 1984. Extrinsic barriers to gene flow in rock-dwelling cichlids of Lake Malawi: macrohabitat heterogeneity and reef colonization. In: Echelle AA, Kornfield I, eds. *Evolution of Fish Species Flocks*. Orono, Maine: University of Maine Press. p. 169-183

McKaye KR, Howard JH, Stauffer JR, Morgan RP, Shonhiwa F. 1993. Sexual selection and genetic relationships of a sibling species complex of bower building cichlids in Lake Malawi, Africa. *Japan. J. Ichthyol.* 40:15-21

McKaye K, Kocher TD, Reinthal P, Harrison H, Kornfield I. 1984. Genetic evidence for allopatric and sympatric differentiation among color morphs of a Lake Malawi cichlid fish. *Evolution* 38:215-219

McKaye K, Kocher T, Reinthal P, Kornfield I. 1982. A sympatric sibling species complex of *Petrotilapia* Trewavas from Lake Malawi analyzed by enzyme electrophoresis (Pisces: Cichlidae). *Zool. J. Linn. Soc.* 76:91-96

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

Meyer A. 1993. Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends Ecol. Evol.* 8:279-285

Meyer A, Knowles LL, Verheyen E. 1996. Widespread geographical distribution of mitochondrial haplotypes in rock-dwelling cichlid fishes from Lake Tanganyika. *Mol. Ecol.* 5:341-350

Meyer A, Kocher T, Basasibwaki P, Wilson A. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* 347:550-553

Meyer A, Ontero C, Spreinat A. 1994. Evolutionary history of the cichlid fish species flocks of the East African great lakes inferred from molecular phylogenetic data. *Arch. Hydrobiol. Beth. Limnol.* 44:407-423

Meyer M, Foerster W. 1984 Un nouveau *Pseudotropheus* du lac Malawi avec des remarques sur le complexe Pseudotropheus-Melanochromis (Pisces, Perciformes, Cichlidae). Revue Française d'Aquariologie et de Herpetologie 10:107-112

Moller AP, Cuervo JJ. 1998. Speciation and feather ornamentation in birds. *Evolution* 52:859-869

Moran P, Kornfield I. 1993. Retention of an ancestral polymorphism in the *mbuna* species flock (Pisces: Cichlidae) of Lake Malawi. *Mol. Biol. Evol.* 10:1015-1029

Moran P, Kornfield I. 1995. Were population bottlenecks associated with the radiation of the *mbuna* species flock of Lake Malawi? *Mol. Biol. Evol.* 12:1085-1093

Moran P, Kornfield I, Reinthal P. 1994. Molecular systematics and radiation of the haplochromine cichlids (Teleostei: Perciformes) of Lake Malawi. *Copeia* 1994:274-288

Nagl S, Tichy H, Mayer WE, Takahata N, Klein J. 1998. Persistence of neutral polymorphisms in Lake Victoria cichlid fish. *Proc. Natl. Acad. Sci. USA* 95:14238-14243

Nei, M. 1972. Genetic distances between populations. *Amer. Nat.* 106:283-292

Nishida M. 1991. Lake Tanganyika as an evolutionary reservoir of old lineages of East African cichlid fishes: inferences from allozyme data. *Experientia* 47:974-979

Nishida M. 1997. Phylogenetic relationships and evolution of Tanganyikan cichlids: a molecular perspective. In: Kawanabe H, Hori M, Naagoshi M, eds. *Fish Communities in Lake Tanganyika*. Sakyo-ku, Japan: Kyoto University Press. p. 1-23

Ogutu-Ohwayo R. 1993. The effects of predation by Nile Perch, *Lates niloticus* L., on the fish of Lake Nabugabo, with suggestions for conservation of endangered endemic cichlids. *Conserv. Biol.* 7:701-711

Ono H, O'hUigín C, Tichy H, Klein J. 1993. Major-histocompatibility-complex variation in two species of cichlid fishes from Lake Malawi. *Mol. Biol. Evol.* 10:1060-1072

Owen RB, Crossley R, Johnson TC, Tweddle D, Kornfield I, et al. 1990. Major low levels of Lake Malawi and their implications for speciation rates in cichlid fishes. *Proc. R. Soc. Lond. B* 240:519-533

Owens I, Bennet P, Harvey P. 1999. Species richness among birds: body size, life history, sexual selection, or ecology? *Proc. R. Soc. Lond. B* 266:933-939

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C. 1997 An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* 147:1943-1957

Page RDM, Holmes EC. 1998. *Molecular Evolution*. Oxford: Blackwell Science. 346

Park SDE. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection [Ph.D. thesis (in prep.)], University of Dublin.

Parker A, Kornfield I. 1996. Polygynandry in *Pseudotropheus zebra,* a cichlid fish from Lake Malawi. *Env. Biol. Fishes* 47:345-352

Parker A, Kornfield I. 1997. Evolution of the mitochondrial DNA control region in the *mbuna* (Cichlidae) species flock of Lake Malawi, East Africa. *J. Mol. Evol.* 45:70-83

Patterson HEH. 1985. The recognition concept of species. In: Vrba ES, ed. *Species and Speciation*. Pretoria: Transvaal Museum. pp. 21-29

Petren K, Grant B, Grant P. 1999. A phylogeny of Darwin's Finches based on microsatellite DNA length variation. *Proc. R. Soc. Lond. B* 266:321-329

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

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): Population genetics software for exact test and ecumenicism. *J. Heredity* 86:248-249

Reinthal P, Meyer A. 1997. Molecular phylogenetic tests of speciation models in African cichlid fishes. In: Givinish T, Sytsma K, eds. *Molecular Phylogenetics of Adaptive Radiations*. Boston: Cambridge University Press. p. 375-390

Reneau P. 2001. Inheritance of dorsal fin coloration in the *Metriaclima* species complex (Teleostei: Cichlidae) of Lake Malawi. unpublished Masters degree thesis. Western Kentucky University. p. 33

Ribbink AJ. 1986. The species concept, sibling species and speciation. *Ann. Mus. Roy. Afr. Centr.* 251:109-116

Ribbink AJ. 1994. Alternative perspectives on some controversial aspects of cichlid fish speciation. *Arch. Hydrobiol. Limnol.* 44:101-125

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.* 18:149-310

Ricklefs RE, Schluter D. 1993. Species diversity: regional and historical influences. In: eds. Ricklefs RE, Schluter D. *Species Diversity in Ecological Communities.* Chicago: University of Chicago Press. p. 350-363

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

Roberts TR. 1975. Geographical distribution of African freshwater fishes. *Zool. J. Linn. Soc.* 57:249-319

Rossiter A. 1995. The cichlid fish assemblages of Lake Tanganyika: ecology, behaviour and evolution of its species flocks. In: Begon M, Fitter AH, eds. *Advances in Ecological Research.* Boston: Academic Press. p. 193-217

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

Rundle H, Schluter D. 1998. Reinforcement of stickleback mate preferences: sympatry breeds contempt. *Evolution* 52:200-208

Ruzzante D. 1998. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Can. J. Fish Aquat. Sci.* 55:1-14

Sage D, Selander RK. 1975. Trophic radiation through polymorphism in cichlid fishes. *Proc. Nat. Acad. Sci. USA* 72:4669-4673

Salzburger W, Baric S, Sturmbauer C. 2002. Speciation via introgressive hybridization in African Cichlids? *Mol. Ecol.* 11:619-625

Schilthuizen M, Hoekstra RF, Gittenberger E. 2001. The 'rare allele phenomenon' in a ribosomal spacer. *Mol. Ecol.* 10:1341-1345

Schliwewen U, Tautz D, Paabo S. 1994. Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368:629-632

Schluter D. 1998. Ecological Causes of Speciation. In: Howard DH, Berlocher SH, eds. *Endless Forms Species and Speciation*. New York: Oxford University Press. p. 114-129

Schneider S, Roessli D, Excofier L. 2000. Arlequin: A software for population genetics data analysis. Geneva: Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Seegers L, Sonnenberg R, Yamamoto R. 1999. Molecular analysis of the *Alcolapia* flock from lakes Natron and Magadi, Tanzania and Kenya (Teleostei: Cichlidae), and implications for their systematics and evolution. *Ichthyol. Explor. Freshwaters* 10:175-199

Seegers L, Tichy H. 1999. The *Oreochromis alcalicus* flock (Teleostei: Cichlidae) from lakes Natron and Magadi, Tanzania and Kenya, with descriptions of two new species. *Ichthyol. Explor. Freshwaters* 10:97-146

Seehausen O. 1999. Explosive speciation rates and unusual species richness in Haplochromine Cichlids - effects of sexual selection. *Adv. Ecol. Res.* 30:325-371

Seehausen O. 2002. Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14,600 year history of its cichlid species flock. *Proc. R. Soc. Lond. B* 269: 491-497

Seehausen O, Lippitsch E, Bouton N, Zwennes H. 1998. Mbipi, the rock-dwelling cichlids of Lake Victoria: description of three new genera and fifteen new species (Teleostei). *Ichthyol. Explor. Freshwaters* 9:129-228

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

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

Seehausen O, Van Alphen JJA. 1999. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.* 2:262-271

Seehausen O, Van Alphen JJA, Lande R. 1999. Color polymorphism and sex ratio distortion in a cichlid fish as an insipient stage in sympatric speciation by sexual selection. *Ecol. Lett.* 2, 367-378

Seehausen O, Van Alphen JJM, Witte F. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277:1808-1811

Seehausen O, Witte F, Van Alphen JJM, Bouton N. 1998. Direct mate choice maintains diversity among sympatric cichlids in Lake Victoria. *J. Fish Biol.* 53:37-55

Snoeks, J. 2000. How well known is the ichthyodiversity of large East African lakes? In: Rossiter A, Kawanabe H, eds. *Ancient Lakes: Biodiversity, Ecology and Evolution*, vol. 31. Boston: Academic Press. p. 17 - 38

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 a description of a new genus and ten new species. *Proc. Acad. Nat. Sci. Phil.* 148:189-230

Stauffer JR, Bowers NJ, Kocher TD, McKaye KR. 1996. Evidence of hybridization between *Cyanotilapia afra* and *Psuedotropheus zebra* (Teleostei: Cichlidae) following an intralacustrine translocation in Lake Malawi. *Copiea* 1996:203-208

Stauffer JRJ, Hert E. 1992. *Pseudotropheus callainos* a new species of *mbuna* (Cichlidae), with analyses of changes associated with two intra-lucustrine transplantations in Lake Malawi, Africa. *Ichthyol. Explor. Freshwaters* 3:253-264

Stiassny MLJ. 1981. Phylogenetic versus convergent relationship between piscivorous cichlid fishes from Lakes Malawi and Tanganyika. *Bull. Brit. Mus. Nat. Hist. Zool.* 40:167-201

Stiassny MLJ, Meyer A. 1999. Cichlids of the rift lakes. *Sci. Am.* 280:44-49

Strickberger MW. 2000. *Evolution*. Sudbury, MA: Jones and Bartlett. 722

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, Ruber L, Verheyen E. 2001. Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Mol. Biol. Evol.* 18:144-154

Sturmbauer C, Meyer A. 1992. Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature* 358:578-581

Sturmbauer C, Meyer A. 1993. Mitochondrial phylogeny of the endemic mouthbrooding lineages of cichlid fishes from Lake Tanganyika, East Africa. *Mol. Biol. Evol.* 10:751-768

Sturmbauer C, Verheyen E, Meyer A. 1994. Mitochondrial phylogeny of the Lamprologini, the major substrate spawning lineage of cichlid fishes from Lake Tanganyika in Eastern Africa. *Mol. Biol. Evol.* 11:691-703

Sturmbauer C, Verheyen E, Ruber L, Meyer A. 1997. Phylogeographic patterns in populations of cichlid fishes from rock habitats in Lake Tanganyika. In: Kocher TD, Stepien CA, eds. *Molecular Systematics of Fishes.* San Diego: Academic Press. p. 97-111

Sültmann H, Mayer WE. 1997. Reconstruction of cichlid fish phylogeny using nuclear DNA markers. In: Kocher TD, Stepien CA, eds. *Molecular Systematics of Fishes.* San Diego: Academic Press. p. 39-51

Takahashi K, Terai Y, Nishida M, Okada N. 1998. A novel family of short interspersed elements (SINEs) from cichlids: the patterns of insertion of SINEs at orthologous loci support the proposed monophyly of four major groups of cichlid fishes in Lake Tanganyika. *Mol. Biol. Evol.* 15:391-407

Takahashi K, Terai Y, Nishida M, Okada N. 2001. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. *Mol. Biol. Evol.* 18: 2057-2066

Takezaki N, Nei M. 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144:389-399

Temple PH. 1969. Some biological implications of a revised geological history for Lake Victoria. *Biol. J. Linn. Soc.* 1:363-371

Templeton AR. 1989. The meaning of species and speciation: a genetic perspective. In: Otte D, Endler JA, eds. *Speciation and its Consequences*. Sunderland, Massachusetts: Sinauer Associates. p. 3-27

Trewavas E. 1935. A synopsis of the cichlid fishes of Lake Nyasa. *Ann. Mag. Nat. Hist.* 10:65-118

Turner GF. 1994. Speciation mechanisms in Lake Malawi cichlids: a critical review. *Arch. Hydrobiol. Beih. Limnol.* 44:139-160

Turner GF. 2000. The nature of species in ancient lakes: perspectives from the fishes of Lake Malawi. In: Rossiter A, Kawanabe H, eds. *Ancient Lakes: Biodiversity, Ecology and Evolution, vol. 31*. Boston: Academic Press. p. 39-60

Turner GF, Burrows MT. 1995. A model of sympatric speciation by sexual selection. *Proc. R. Soc. Lond. B* 260:287-292

Turner GF, Seehausen O, Knight ME, Allender CJ, Robinson RL. 2001. How many species of cichlid fishes are there in African lakes? *Mol. Ecol.* 10:793-806

Van Dorn GS, Noest AJ, Hogeweg P. 1998. Sympatric speciation and extinction driven by environment dependent sexual selection. *Proc. R. Soc. Lond. B* 265:1915-1919

Van Oppen MJH, Rico C, Deutsch JC, Turner GF, Hewitt GM. 1997a. Isolation and characterization of microsatellite loci in the cichlid fish *Pseudotropheus zebra*. *Mol. Ecol.* 6:387-388

Van Oppen MJH, Rico C, Turner GF, Hewitt GM. 2000. Extensive homoplasy, nonstepwise mutations and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi Cichlids. *Mol. Biol. and Evol.* 17:489-498

Van Oppen MJH, Turner GF, Rico G, Deutsch JC, Ibrahim KM, Robinson, RL, Hewitt GM. 1997b. Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc. R. Soc. Lond. B* 264:1803-1812

Van Oppen MJH, Turner GF, Rico C, Robinson R, Deutch J, et al. 1998. Asortative mating among rock-dwelling cichlid fishes supports high estimates of species richness from Lake Malawi. *Mol. Ecol.* 7:991-1001

Verschuren D, Johnson TC, Kling HJ, Edgington DN, Leavitt PR, Brown ET, Talbot MR, Hecky RE. 2002. History and timing of human impact on Lake Victoria, East Africa. *Proc. R.. Soc. Lond.. B* 269:289-294

Witte F, Barel KDN, Van Oijen MJP. 1997. Intraspecific variation of haplochromine cichlids from Lake Victoria and its taxonomic implications. *S. Afr. J. Sci.* 93:585-594

Wu L, Kaufman L, Fuerst P. 1999. Isolation of microsatellite markers in *Astatoreochromis alluaudi* and their cross-species amplifications in other African cichlids. *Mol. Ecol.* 8:895-906

BIOGRAPHY OF THE AUTHOR

Peter Smith was born in Hartford, Connecticut in 1974, one of the six children of Mary Ellen and David Smith. He graduated from Hall High School in West Hartford in 1993, and later earned a B.A. in Zoology from the University of Maine, in 1997. He is married to Ashley Begin-Smith and is presently working as a postdoctoral fellow at Maine Medical Center Research Institute in Scarborough, Maine. Current research interests are in the areas of evolution and development, particularly of vertebrate animals. Peter is a candidate for the Doctor of Philosophy degree in Biological Sciences from the University of Maine in December, 2002.