DIVERSIFICATION IN AFRICAN CICHLID FISHES: FROM SPECIATION TO MACROEVOLUTIONARY PATTERNS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Catherine E. Wagner

August 2011

DIVERSIFICATION IN AFRICAN CICHLID FISHES: FROM SPECIATION TO MACROEVOLUTIONARY PATTERNS

Catherine E. Wagner, Ph. D.

Cornell University 2011

African cichlid fishes are a celebrated example of evolutionary diversification, and a thorough understanding of this diversification requires studies of both the origins of their diversity, and of the factors influencing broad scale evolutionary patterns. This body of work addresses evolutionary questions at these both of these levels. I first test the influence of habitat and distance on fine-scale population genetic structure in three sympatric cichlid species in Lake Tanganyika. These species show striking differences in their patterns of genetic subdivision within the same geographical region, implying substantially different patterns of gene flow. This suggests that both ecological and behavioral traits have a strong influence on the scale and degree of population subdivision, a finding which has implications for understanding differential propensities for diversification among lineages. Next, I study the population genetics of recent speciation in sympatric color morphs from Lake Tanganyika. I report genetic evidence that these color morphs diverged only recently, yet that barriers to gene flow exist which prevent extensive admixture despite their sympatric distribution. This is an unusual example of active diversification in Lake Tanganyika's ancient cichlid fauna. Moving to a macroevolutionary perspective, I next examine the factors that influence cichlid adaptive radiation. Cichlids have radiated within more than thirty African lakes, and in another seventy instances, colonizing lineages are present in lakes without diversifying. Using this "natural experiment", and a dataset

including both environmental variables and information about the traits of colonizing cichlid lineages, I find that lineage-specific traits related to sexual selection, and environmental factors related to ecological opportunity, both strongly influence whether cichlids radiate. Finally, I examine the environmental influences on the species richness of cichlids within African lakes. I show that total species richness per lake is correlated with measures of lake size and energy, and that the species richness of radiations is limited by these same environmental variables. I conclude that ecological carrying capacities exist which render the total diversity of cichlids predictable within these lakes, but that these diversities are achieved by lineage-specific diversification outcomes, thereby producing marked differences in the faunal composition of different lakes.

BIOGRAPHICAL SKETCH

Catherine E. Wagner was born in Berkeley, California in February 1982. From her earliest years, her parents encouraged her to view the world through a scientific lens, and to value the natural world. Her kinship with other species became apparent early on, and among her happiest childhood memories are close encounters with the animals in Tilden Park, the Oakland Zoo, and the Monterey Bay Aquarium.

In 1987 Katie's family moved to Seattle, Washington, where she came to love the dramatic landscapes of Puget Sound and the Pacific Northwest. Her first experience as an ichthyologist came when she took the job of chief salmon caretaker in a third grade class project to raise eggs from a nearby hatchery. She spent her happiest recess hours carefully counting the number of eggs and fry in the classroom's tank and recording these numbers in a chart on the wall. Throughout her childhood and teen years, much of Katie's time outside school was spent with horses. It was on horseback, exploring trails in the northwest forest, that she realized her love of exploring the outdoors in the company of animals.

In the fall of 2000, Katie began her undergraduate education at Whitman College in Walla Walla, Washington. She soon developed strong interests in the history of the natural world, both living and geological, and in examining the interaction between organisms and the landscapes they inhabit as both change over time. An interest in paleoanthropology took her for a summer to a field school in South Africa, where she had her first taste of Africa and began to envision future research in evolutionary biology. In 2003 she spent the summer in Kigoma, Tanzania, on the shores of Lake Tanganyika, with the NSF REU Nyanza project, doing research on the lake's endemic snails under the guidance of Ellinor Michel. Lake Tanganyika is a truly magical place

for a budding evolutionary biologist, and it was here that she resolved to do graduate research in evolutionary biology. Katie graduated *magna cum laude* from Whitman in 2004 with an honors B.A. in Biology and Geology.

In the fall of 2004 Katie enrolled as a Ph.D. student at Cornell University, where she quickly realized that the opportunity to continue research in Lake Tanganyika through the infrastructure of the continuing Nyanza Project was too good to pass up. Her advisor, Amy McCune, soon instilled in her a great appreciation for the evolutionary biology of fishes, and the remarkable cichlids of Lake Tanganyika became a clear choice for her dissertation work. She did not, however, leave behind her initial Tanganyikan interests, and continues her work on population genetics, speciation and diversification of Tanganyika's endemic snails with Ellinor Michel. During her time as a Ph.D. student at Cornell, she had the unusual opportunity to spend a year as a postdoc working in Ole Seehausen's group at EAWAG in Switzerland, and thus added research on macroevolutionary patterns in lakes across Africa to her prior dissertation work on the population genetics and phylogenetics of Tanganyika's cichlids.

In grad school, Katie cultivated a particular interest in teaching and in promoting science education among young people, through many semesters as a teaching assistant and participation as an educator in the annual "Expanding Your Horizons" conference held at Cornell. She was honored to receive the CALS Outstanding Teaching Award in 2010.

In the fall of 2011, Katie will return to the Seehausen group to continue her research on African cichlids. In her free time, Katie plans to explore Switzerland's alpine valleys, lakes and peaks with her trusty canine companion Kiwi.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the help, advice and support of countless people. Amy McCune has been an outstanding advisor, teacher, role model, and source of support throughout my time at Cornell. In her courses, her clarity and enthusiasm have been inspiring to me, both as a student and as her TA. Her pep talks when times were rough, and her unflagging encouragement to carry on, kept me going in both good times and difficult ones. I owe many thanks to the rest of my committee as well. Irby Lovette was a tremendous source of encouragement for me starting with the first grant proposal I wrote, and continuing through my final years at Cornell. He maintains a remarkable awareness of goings-on in the lives of his students, and this awareness, and the support and encouragement that followed from it, helped me immeasurably during the process of completing my Ph.D.. My work has greatly benefitted from Rick Harrison's in-depth knowledge of the biology of speciation, and I have much appreciated the dose of healthy skepticism that he brought to my committee meetings. Ellinor Michel gave me my start in Lake Tanganyika, and made possible my continued work there during my Ph.D.. I am tremendously grateful to her for giving me this opportunity, and for her encouragement of my continued work there. It has been a truly life-changing experience. Ellinor's many years of experience studying Lake Tanganyika and its evolution have also contributed enormously to my work.

The McCune lab group has been a great source of ideas and inspiration; thanks for this to Amy, Mandy Cass, Ezra Lencer, and Dan Rabosky, and many other members over the years. I owe special thanks to Dan Rabosky, not only for being my number-one person to bounce ideas off of for many years, but also for introducing me to R, which has changed the course of my research in

profound ways.

My fieldwork in Kigoma, Tanzania was supported by the large and wonderful staff of the Nyanza Project. First and foremost among them for me was George Kazumbe, who was a crucial component to the success of this dissertation. Never a finer fisherman has been sighted on the shores of Tanganyika as George Kazumbe. I learned from George every field day I spent with him, and his encyclopedic knowledge of Tanganyikan cichlids is remarkable; it helped me in countless ways. George's ability to keep laughing and smiling despite the challenges of the day made even the most difficult of field days fun. I owe so much to him.

Jon Todd was in Kigoma during my field work in both 2005 and 2007, and his unstoppable enthusiasm for snails and macroevolution was not only fun to be around, but it was also a great motivation to keep the ultimate questions in mind during times of field work drudgery. Our collaborations on the diversification of Tanganyikan snails are ongoing, and I look forward to many more years of collaboration with both Ellinor and Jon. I also thank Ellinor and Jon for their hospitality during my visits to London over the years; spending time in the Natural History Museum has always been inspiring.

Although I had heard much of the Great Pete McIntyre before starting at Cornell, it was with tremendous fortune that I happened to become his officemate upon my arrival in Corson Hall. Sharing an office for those first two years had a huge impact on the work that became my dissertation. Pete was the person who encouraged me to work on *Petrochromis*, and was foundational to the design of the study that became Chapter 1 of this dissertation. He also pointed out the interesting *P.* sp. "kazumbe"–*P. polyodon* color variation in the Kigoma region to

me, prior to my first field season collecting fish; this became the subject of Chapter 2 of this dissertation. Pete's guidance has been fundamental to the success of my work, and our collaborations have expanded my thinking in varied and important ways.

I completed the genetic work for my dissertation in Irby Lovette's lab, the Fuller Evolutionary Biology Program at the Cornell Lab of Ornithology. The Lab of O has felt like a second home for me throughout my time at Cornell, and I thank all of Irby's lab group for welcoming me there (despite the paucity of my knowledge of birds). It has been a wonderful community to be part of, and my work has benefitted tremendously from being part of the group. Thanks especially to Laura Stenzler, Matt Carling, Amanda Talaba, Chris Makarewich, Doug Morin, Aurélie Coulon, Rachel Vallender, Mari Kimura, Becky Cramer, Roi Dor, Andrea Townsend, and Elise Ferree.

I have many other colleagues to thank for discussions and ideas over the years. My officemates have been great resources and friends; thanks to Kurt Galbreath, Quinn Thomas, and Gui Becker. Thanks to Jeanne Robertson, Lauren Chan, Nancy Chen, Petra Deane, Erica Larson, Findley Ransler, Marie Nydam, Ben Freeman, Marjorie Weber and Ezra Lencer for discussions of population genetics, phylogenetics, speciation, and comparative methods.

Ole Seehausen and Luke Harmon are my collaborators on chapters 3 and 4 of this dissertation. I have learned much from working with both of them over the past two years, and I look forward to our continued collaborations. My year in Switzerland was also successful thanks to many colleagues and friends. Special thanks go to Martine Maan and Blake Matthews for discussions that filled me with excitement about work to come; I look forward to many more of these brainstorming sessions overlooking Pilatus and Lake Lucerne. Many thanks also to KellyAnn Ross, Julian Junker, Jeff Carpenter, Oliver Selz, and Dörte Carstens for helping me to navigate

the quirky ins and outs of Swiss society and for being my companions on many alpine adventures.

Many sources of funding supported this work. Field work in Lake Tanganyika and the genetic work associated with it was funded by the Cornell University Mellon Fund, the Cornell Department of Ecology and Evolutionary Biology, the Cornell University Graduate School, the Society for Systematic Biology Research Award, the American Society for Ichthyology and Herpetology Raney Fund, the American Cichlid Association Jordan Endowment Fund, Sigma Xi, and an NSF Doctoral Dissertation Enhancement Program Award (INT-0724247). Work associated with Chapters 3 and 4 was supported by the Swiss National Science Foundation and EAWAG, the Swiss National Institute for Aquatic Science and Technology.

I deeply thank the friends who have supported me throughout my years in Ithaca. Dena Vallano and Marie Nydam have been unfaltering friends from my first days at Cornell. Special thanks also to Marita Davison, Jen Moslemi, Elise Ferree. I feel lucky to have had Danica Lombardozzi as part of my life since we were roommates together in Kigoma in 2003; it's not frequent in these oft-uprooted lives we lead that eight years later, we'd still be close friends. I especially thank her for her support during my final year at Cornell. Thanks also to Danica, Jon DeCoste, and the rest of the "south hillians" for adopting me (and Kiwi) into their neighborhood this year. Marjorie Weber has become a wonderful friend in my final year here; I only wish our years at Cornell could have overlapped more. Thanks to Marjorie and Chris Bliss for cheering me up in the times I needed it most–they are the kind of friends you wish you could carry with you wherever you go. Thanks to Ben Dalziel, who brought much joy to my final semester at Cornell, and patiently listened to many hours of mutterings about cichlids. His reassurance and company have made the

lonely work of finishing a Ph.D. a much happier time, and I am overjoyed that I can now be out sailing instead of waving at the boat from the window. I hope that the future will bring many more adventures our way.

I would not be here without the constant love and support from my family; I have everything to thank them for. Thanks to Mary and Madison Wright for being my “local” family all these years. Having them here made Ithaca seem like a familiar place from my first day in town. Thanks to Mia the cat, who infused joy into my life at a time when it was sorely needed, and to Kiwi the dog for being the most dependable of friends, and one who can always make me laugh. Thanks to my sister, Lisi, and my brother, Patrick; they have both had a tremendous impact on the person I have grown up to be. And many, many thanks to my parents, who encouraged me to think about the world as a scientist from the day I started asking questions about the world around me. I can’t express how much their constant love and encouragement has meant to me. Thus far this path has only taken me farther from home, but I hope that some day it will bring me back.

Final thanks to all of the beautiful fish who gave their lives to make this work possible.

TABLE OF CONTENTS

CHAPTER 1

CONTRASTING PATTERNS OF POPULATION GENETIC STRUCTURE IN SYMPATRIC ROCK-DWELLING CICHLID FISHES[1]

Abstract

The cichlid fishes of Lake Tanganyika in Eastern Africa are a celebrated example of both ecological and species diversification. Because population subdivision is likely to play an important role in the speciation process, understanding how habitat features interact with species' demographic, behavioral and ecological attributes to influence gene flow and population divergence may help explain the causes of high species richness in this and other systems. Here, we test the roles of isolation-by-habitat and isolation-by-distance in generating fine-scale population genetic structure in three sympatric species of habitat-restricted cichlids in Lake Tanganyika. Using multi-locus microsatellite genotypes, we contrast patterns of population differentiation in these habitat specialists along a mosaic coastline of both favorable and unfavorable habitat. Despite their close phylogenetic relationship and shared habitat affinity, these species show striking differences in their pattern of genetic subdivision within the same geographical region, suggesting substantially different patterns of gene flow. In particular, two trophically specialized species exhibit much more restricted gene flow over sandy habitat than a trophically opportunistic species. This result suggests that ecological and behavioral traits have a strong influence on the scale and degree of population subdivision, a finding which has

[1] This chapter has been published as Wagner, C. E., and A. R. McCune. 2009. Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. Evolution 63:1312-1326,; it is reprinted here with permission.

potentially important implications for understanding differential propensities for diversification among lineages and phylogenetic patterns of diversity.

Introduction

Gene flow is a fundamental process in the evolution of populations, and the central influence on population genetic structure. At one extreme, extensive gene flow may lead to genetic homogenization of populations. At the other extreme, little or no gene flow may permit the evolution of reproductive barriers between populations, with or without adaptive divergence (Schluter 2000; Rundle and Nosil 2005). Because the spatial genetic structure of populations often plays a key role in speciation (Gavrilets 2003; Coyne and Orr 2004), understanding the factors that promote genetic isolation between populations provides important context for the process of speciation.

Extrinsic habitat features and intrinsic organismal traits synergistically influence levels of gene flow among populations. The dispersal of individuals and the distribution of populations in a landscape are influenced by both species' traits and by the heterogeneity of the landscape itself. A habitat feature that strongly deters gene flow in one species may not have the same effect in another species; the strength of a barrier depends critically on the life history, ecology, and behavior of an organism. Although a species' ecology may affect speciation through its influence on the distribution of populations (Wiens 2004) and dispersal (e.g. Smith and Farrell 2006), most recent work on the role of ecological specialization in diversification has been largely devoid of discussion of specialization's impact on population genetic structure (Schluter 2000; e.g. Ackermann and Doebeli 2004). However, if a lineage's ecological traits have a repeated and

consistent influence upon gene flow among populations, these traits may have a profound impact on the lineage's ability to diversify (Vrba 1984).

Studying multiple species across a shared landscape allows inference of the similarities and differences in species' traits that influence their population genetic structure, and provides insight into habitat factors that impact genetic structure across species (Castric and Bernatchez 2004; Whiteley et al. 2004; Manier and Arnold 2006). Furthermore, this approach can reveal differences in the relative scales of population genetic structure among species. Links between population genetic structure and propensity to speciate have been hypothesized, with the reasoning that greater population genetic structure increases opportunities for genetic divergence among populations, and thereby affords greater opportunity for speciation (e.g. Rice 2004; Vrba and Gould 1986). By this argument, if population genetic structure consistently influences speciation, species from highly diverse groups would be expected to exhibit stronger patterns of population genetic structure than species from low-diversity groups. Here we apply a comparative population genetic approach to examine the factors shaping population genetic structure, and their potential influence on diversification, in three cichlid species from Lake Tanganyika.

Background on cichlid speciation.

The exceptionally high species diversity, speciation rates, and degree of morphological and ecological diversity among African cichlids makes them a model for the study of explosive diversification (Kocher 2004; Seehausen 2006). Cichlids and other endemic fish species in freshwater lakes have attracted considerable attention as potential cases of sympatric speciation (Schliewen et al. 1994; Schluter 1996; e.g. Barluenga and Meyer 2004) because of the

assumption that within single bodies of water, mobile fish will not encounter physical barriers to gene flow (Coyne and Orr 2004). However, there is now considerable evidence that lake fish populations can exhibit genetic differentiation between populations at fine geographic scales (Barluenga and Meyer 2005; Adams et al. 2006; Bergek and Bjorklund 2007). Although cases of speciation in small lakes remain among the best-supported examples of sympatric speciation (e.g. Schliewen et al. 1994), in lakes large enough that cryptic barriers to dispersal may exist, allopatric or parapatric speciation scenarios should be investigated, especially given the limited theoretical conditions under which sympatric speciation may occur (e.g. Arnegard and Kondrashov 2004; Bolnick 2004; Burger and Schneider 2006).

Classic models for cichlid speciation involve the allopatric differentiation of species through subdivision of populations in heterogeneous littoral habitat (Trewavas 1947; Fryer 1959; Fryer and Iles 1972). Fluctuations in lake level would, in these models, cause populations to fuse and fragment, influencing the process of population divergence at varying spatial and temporal scales dependent on the severity and tempo of lake level fluctuation (Sturmbauer and Meyer 1992; Verheyen et al. 1996). An extremely old and deep basin (9-12 million years; Cohen et al. 1993), Lake Tanganyika has been subject to an extended history of lake level fluctuation, with historical drops in lake level at least once subdividing the lake into three separate basins (Tiercelin and Mondeguer 1991). Although many studies in Lake Tanganyika have focused on large-scale phylogeographic patterns (Sturmbauer and Meyer 1992; Meyer et al. 1996; Verheyen et al. 1996; e.g. ~450 km - whole-lake: Baric et al. 2003; Sturmbauer et al. 2005; Egger et al. 2007), few have focused on population genetic divergence at relatively fine geographic scales (Taylor et al. 2001; e.g. 40-60 km: Duftner et al. 2006; Koblmuller et al. 2007; Sefc et al. 2007). Although

large-scale vicariant events have dramatic impacts on divergence because of their ability to definitively halt gene flow, major lake subdivisions are historically rare in comparison to ongoing smaller-scale fluctuations throughout Tanganyika's and Malawi's histories, and vicariant events are absent in Lake Victoria's history (Cohen et al. 2007). If divergence between geographically close populations accumulates quickly due to restricted gene flow at fine geographic scales, ongoing minor lake level fluctuations will provide opportunities for diverging populations to be challenged episodically by sympatry. Understanding the geographic scales at which population genetic divergence accumulates, and the habitat and ecological factors influencing this divergence, is crucial for understanding the origination of cichlid species in rift lake habitats.

For the diverse cichlids of the East African rift lakes, ecological ties to rocky substrate might be influential in determining the degree to which populations are structured within a heterogeneous habitat. Several studies of Lake Malawi's diverse rock-dwelling "mbuna" cichlids have found significant differentiation across sandy bays and stretches of deep water (Van Oppen et al. 1997; Arnegard et al. 1999; Markert et al. 1999; Danley et al. 2000). In Lake Tanganyika, rock-dwelling Eretmodine cichlids exhibit significantly lower gene flow over sandy substrate, but also show strong isolation by distance over stretches of rocky substrate (Taylor et al. 2001). Four other Tanganyikan cichlid species have been shown to exhibit substantial reduction in dispersal across a single sandy bay in the south of Lake Tanganyika (Koblmuller et al. 2007; Sefc et al. 2007). Evidence from a variety of cichlid species thus suggests that the distribution of habitat features can shape population genetic patterns. However, traits underlying the strength of a species' ecological ties to habitat features, in this case rocky substrate, will also shape population

genetic structure. For example, those species more capable of using diverse food resources may be expected to disperse more across stretches of inappropriate habitat than those incapable of dietary flexibility. Likewise, species that feed in the water column over rocks may be more prone to move across inappropriate habitat than do species that utilize benthic habitat. Thus, by influencing population genetic structure, ecological traits may play an important role in speciation and patterns of diversification.

Study system and design.

This study examines population genetic structure in three sympatric species of rock-dwelling cichlids from Lake Tanganyika. All study species are closely related members of the tribe Tropheini: *Petrochromis* sp. *"*kazumbe*"*, *P.* sp. *"*moshi*"* and *Simochromis diagramma* (Figure 1.1). Both *Simochromis* and *Petrochromis* species primarily eat algae off of rocks: *Petrochromis* species are specialized algivores that graze on epilithic algae, *S. diagramma* is primarily a browser, eating filamentous algae off of rock surfaces (Yamaoka 1983). However, whereas *S. diagramma* is capable of opportunistic feeding (Yamaoka, 1983; C. Wagner, personal observations), the highly specialized morphology of *Petrochromis* species substantially limits trophic flexibility (Yamaoka 1983). All three species are polygamous, and maternally mouth-brood their young (Brichard 1989). They are sympatric over rocky substrates in depths up to approximately 10 meters in the Kigoma region of Tanzania (Figure 1.2). Although both genera are in need of taxonomic revision and thorough phylogenetic analysis, widespread collections from the aquarium trade indicate that *Petrochromis* is substantially more diverse than *Simochromis* (*Petrochromis*: 18 recognized morphs, 6 described as species; *Simochromis*, 7 recognized morphs, 5 described species; Brichard 1989; Herrmann 1996). In the Kigoma region

up to 6 *Petrochromis* morphs occur in sympatry whereas only two *Simochromis* morphs are present (C. Wagner, unpublished data). Tropheines are the endemic Tanganyikan sister clade to the Lake Malawi and the Lake Victoria haplochromine cichlid radiations (Salzburger et al. 2005).

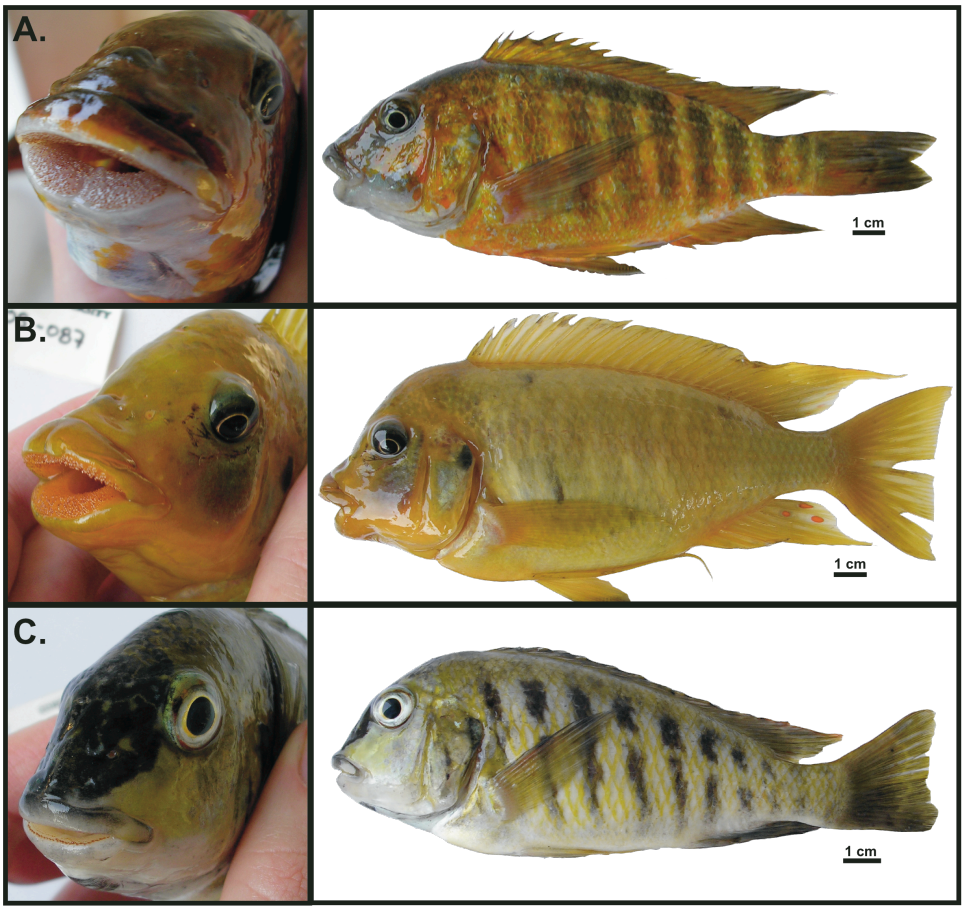


Figure 1.1. Study species: A) *Petrochromis* sp. "kazumbe", B) *P.* sp. "moshi", C) *Simochromis diagramma*.

Populations in this study are from a 60 km stretch of shoreline bordering the region of Kigoma, Tanzania (Figure 1.2). Several hundred meters west of the current Kigoma shoreline is a large basin-bounding fault, causing the shoreline to be extremely steep in most locations (Soreghan and Cohen 1996). Because the lake is anoxic below approximately 65 meters (Coulter 1991), and the depth tolerance of study species is substantially less than this lower bound, these and all other

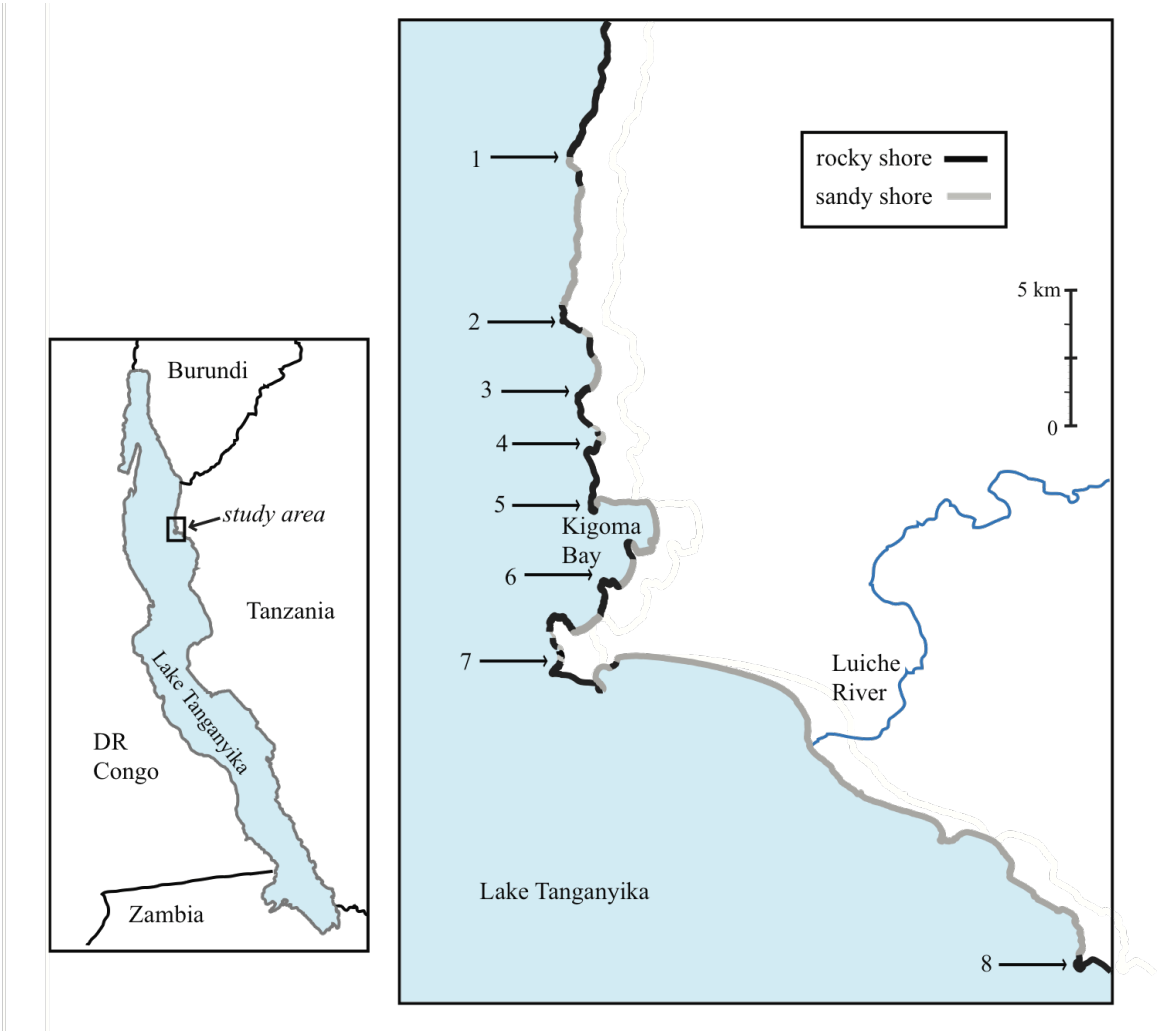


Figure 1.2. Fishes were collected from eight sampling sites on the shoreline of Lake Tanganyikabordering the Kigoma region of Tanzania. Shore with rocky substrate is depicted in black, and sandy shoreline is indicated in grey.

benthic species are restricted to a linear strip of rocky shoreline interspersed with sandy patches of differing sizes. Adjacent populations are therefore expected to exchange more migrants than any non-adjacent site pair. These geographic population limits simulate the classic 'stepping-stone' model in population genetics, and create the potential for strong patterns of genetic isolation by distance (Wright 1943; e.g. Kimura and Weiss 1964; Slatkin 1993). The predominance of rocky substrate in the Kigoma region is due to the exposure of quartzite and

conglomerate strata along the current shoreline. Near-shore sand substrate patches result either from offshore deposition from streams and rivers, or the exposure of sandstone strata at the shoreline (the Manyovu Red Beds; Tiercelin and Mondeguer 1991). In this latter case, because of the more erodible nature of sandstone than quartzite or conglomerate exposures, embayments along the coastline often represent areas of sandy substrate.

Our goal in this study was to examine the factors shaping population genetic structure in three cichlid species from Lake Tanganyika, employing a comparative population genetic approach. We predicted that sand substrate would impede gene flow relative to rock substrate in sympatric *P.* sp. *"*kazumbe*"*, *P.* sp. *"*moshi*"* and *S. diagramma* based on their shared affinity for rocky habitats. In addition, based on their greater capacity for trophic flexibility, we predicted that *S. diagramma* would exhibit less spatial population genetic structure than *Petrochromis* species within the study region. We tested these predictions using multi-locus microsatellite genotypes and Bayesian analysis of population structure, and examined and compared patterns of population subdivision in these three cichlid species. We used data on the distribution of substrate types along the shoreline of the study area to examine the relationship between population structure and habitat type. Understanding the drivers of population genetic structure, and differences in scales of population genetic structure among species, has important implications for understanding cichlid speciation and patterns of diversification.

Methods

Sample collection

Three species were collected for this study: *Petrochromis* sp. "kazumbe", *P.* sp. *"*moshi" and *Simochromis diagramma* (Figure 1.1). Both *Petrochromis* species are currently undescribed,

although they are well-known from the aquarium trade and readily identified based on color pattern and genetic data (C. Wagner, unpublished data). We collected fishes using gill nets from a total of eight sites spanning ~60 km of coastline in the region of Kigoma, Tanzania (Figure 1.2). Fishes from sites 4-8 were collected in 2005 and fishes from sites 1-3, plus additional samples from sites 4-7, were collected in 2007. All sites were sampled over the course of several days of collecting, with the exception of site 8, at which all fishes were collected during one day. *Petrochromis* sp. "moshi" were sufficiently rare at site 2 that they are absent from collections at that site. Fin clips were taken from each fish and preserved in DMSO-EDTA buffer (Seutin and White 1991). All specimens were retained as vouchers and have been deposited in the Cornell Museum of Vertebrates (Appendix 1, Table S1).

DNA extraction, PCR, genotyping

We extracted genomic DNA from fin clips using DNeasy Tissue Kits (Qiagen). We then used polymerase chain reaction (PCR) to amplify 12 microsatellite loci previously developed for East African cichlids (Table 1.1). Forward primers were labeled at the 5' end with fluorescent tags (PET, 6-FAM, VIC or NED; Applied Biosystems), and a 5' tailing sequence (GTTTCT) was added to most reverse primers to force adenylation of PCR products, thereby standardizing allele sizes (Brownstein et al. 1996). Individual PCR reactions were combined into multiplex mixes in order to decrease the number of reactions needed per individual, reducing the necessary PCR reactions per individual from 12 to 5, with up to three loci amplified per reaction. PCR reactions (10ul) included 10-100 ng of genomic DNA, 0.25 U of Jumpstart^{TM} *Taq* polymerase (Sigma), 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1 – 1.5 mM $MgCl_2$ specific to each locus, 200 uM of

dNTPs (Invitrogen) and 1.3 - 3 pmol each of forward and reverse primers, concentration dependent on the strength of the locus' amplification in the multiplexed PCR (Hailer et al. 2005). Cycling profiles for PCR reactions included an initial 4 minutes at 94°C followed by 35 cycles of 30 seconds or 1 minute (depending on reaction) at 94°C, 30 seconds or one minute at 52°C, 56°C, 50°C or 48°C (specific per multiplexed PCR mix), and one minute at 72°C, followed by a final incubation at 72°C for 30 minutes. Products of PCR reactions were genotyped in two genotyping mixes chosen to optimize size separation of loci with the same fluorescent dyes. Allele sizes were estimated using GeneMapper version 3.7 and verified and amended by eye in order to fix miscalled peaks (Applied Biosystems).

Table 1.1. Allelic diversity and size range for the twelve microsatellite loci used in this study for species *P. kazumbe*, *P. moshi*, and *S. diagramma*. Loci were originally described in other East African cichlid species in the references indicated.

	***Petrochromis* sp. "kazumbe" (N=218)**		***Petrochromis* sp. "moshi" (N=155)**		***Simochromis diagramma* (N=210)**		
Locus	**Total alleles**	**size range**	**Total alleles**	**size range**	**Total alleles**	**size range**	**Reference**
TmoM5	44	328-391	26	336-382	38	315-404	Zardoya et al. 1996
TmoM7	34	326-409	10	269-353	11	273-343	Zardoya et al. 1996
TmoM11	1	--	11	169-250	49	173-233	Zardoya et al. 1996
TmoM13	41	217-271	14	210-405	58	203-285	Zardoya et al. 1996
TmoM25	19	359-405	5	337-366	9	344-391	Zardoya et al. 1996
Pzeb1	20	125-161	16	123-161	76	122-233	van Oppen et al. 1997
Pzeb4	3	118-122	5	118-127	20	114-169	van Oppen et al. 1997
Pzeb5	1	--	2	125-127	5	123-136	van Oppen et al. 1997
UNH001	5	125-151	1	--	9	142-232	Kellogg et al. 1995
UNH002	14	178-207	19	182-232	34	172-242	Kellogg et al. 1995
UME002	14	237-282	4	232-245	39	235-323	Parker and Kornfield 1996
UME003	7	161-187	20	173-265	28	173-243	Parker and Kornfield 1996

Data analysis

We tested for significant linkage disequilibrium among all pairs of loci in all populations using GENEPOP v4, and 10000 dememorizations, 500 batches, and 5000 iterations per batch to test for significance by the Markov chain algorithm (Raymond and Rousset 1995). We calculated the

observed, expected and non-biased expected heterozygosity for each population at each locus using GENETIX 4.05.2 (Belkhir et al. 1996-2004). We then tested for significant deviations from Hardy-Weinberg equilibrium using the exact test implemented in GENEPOP v4 (Haldane 1954; Guo and Thompson 1992; Weir 1996), using a Markov chain algorithm to test for significance. We set Markov chain parameters for all tests so that they reduced the standard error of p-values to under 0.01. For tests within populations and within loci, Markov chain parameters were set to 10000 dememorizations, 200 batches and 5000 iterations per batch. We also tested for a global deficiency of heterozygotes across all populations and all loci via the multisample score test implemented in GENEPOP v4 (Raymond and Rousset 1995) using Markov chain parameters of 10000 dememorizations, 1000 batches and 10000 iterations per batch.

To correct for multiple comparisons in all of the applicable analyses, we used false discovery rate (FDR) correction (Benjamini and Hochberg 1995), implemented in the R programming environment (R Development Core Team 2008) using the library of functions available at http://www.stjuderesearch.org/depts/biostats/documents/fdr-library.R (Pounds 2006).

To examine population structure, we first calculated global F_{ST} values and F_{ST} values for all pairs of population samples in GENETIX 4.05.2 (Belkhir et al. 1996-2004), using Weir and Cockerham's (1984) estimator and testing for significance via permutation tests of 10,000 replicates. Because F_{ST} estimators are based on levels of heterozygosity, they are affected strongly by the variability of the genetic marker used. Highly variable markers will result in lower F_{ST} values than lower variability ones given the same level of population structure (Hedrick 2005; Meirmans 2006). In order to produce estimates of population structure that were comparable among study species in spite of potentially different effective population sizes and

different levels of heterozygosity, we used Meirmans' (2006) standardized statistic, φ'_{ST}. This statistic is based on Excoffier et al.'s (1992) φ_{ST}, an AMOVA analog of F_{ST}, but it is equivalent to Weir and Cockerham's (1984) estimator when used as a measure of pairwise population differentiation (Meirmans 2006). Meirmans' (2006) φ'_{ST} is standardized by dividing φ_{ST} by the maximum possible φ_{ST} given the observed within-population variance. This standardization makes the statistic directly comparable among markers of differing diversity, or among species with different effective population sizes (Meirmans 2006). To calculate φ'_{ST}, we first transformed our dataset so as to maximize among-population variance while holding within-population variance constant, and calculated $\varphi_{ST(max)}$ from this transformed data set using GenoDive 2.0b9 (Meirmans and Van Tienderen 2004). We then calculated φ'_{ST} by dividing the actual φ_{ST} value by the theoretical $\varphi_{ST(max)}$ value.

To assess the effects of individual loci on estimates of F_{ST} we performed jackknife analyses over loci for each species. At each round in the jackknife procedure we removed one locus and recalculated pairwise F_{ST} estimates using the package *hierfstat* (Goudet 2005, 2006) in the R programming environment (R Development Core Team 2008). R functions to perform these analyses over loci are available upon request from CEW.

We further analyzed patterns of population structure within each species using the Bayesian clustering program STRUCTURE v.2.2 (Pritchard et al. 2000). STRUCTURE assigns individuals to *K* populations, maximizing Hardy-Weinberg and linkage equilibrium within populations. For each species, we ran STRUCTURE for *K*=1 through *K*=10, with 10 iterations at each *K* value and using 180,000 MCMC (Markov Chain Monte Carlo) generations as burn-in,

followed by 1 million generations to generate the posterior sample distribution. We examined lnP (D), the probability of the data given *K*, over the course of the burn-in and the run to ensure that these values had stabilized by the end of the burn-in period. All runs used the admixture and correlated allele frequency models, and we set the model to infer the parameter alpha, which denotes degree of admixure, separately for each population, under the assumption that levels of admixture were likely to differ substantially among populations. To assess optimal values of *K* from STRUCTURE runs, we first used the method suggested by Pritchard and Wen (2003) to find the most likely value of *K* by comparing lnP(D) values from runs of different *K*. We plotted the lnP(D) values from the posterior distribution of all runs to graphically confirm this result. If this value of *K* was greater than 1, we also used the Δ*K* method of Evanno et al. (2005) to assess the most likely value of *K* for the dataset. Evanno et al.'s (2005) method finds the breakpoint in slope for the distribution of lnP(D) values over the assessed values of K. In situations where there is hierarchical substructure in a dataset, the optimal Δ*K* solution will represent the uppermost level of population substructure (Evanno et al. 2005). Therefore, we implemented this method in a hierarchical framework in order to fully identify substructure in the datasets (see Coulon et al. 2008). We calculated optimal Δ*K* values for each dataset where *K* inferred from Pritchard and Wen's (2003) method was greater than one. Using the program *CLUMPP* (Jakobsson and Rosenberg 2007), we generated consensus percentages of inferred ancestry from the 10 original STRUCTURE runs of this *K* value. Using these consensus percentages of inferred ancestry, we subdivided the original dataset into *K* subsets by assigning individuals to the group corresponding to their highest inferred ancestry percentage. We assigned all individuals whose ancestry was greater than 0.6 to a group; individuals with inferred ancestry of <0.6 for all groups

were not used in further analyses. We then repeated STRUCTURE runs on these subsets using the same run parameters as in the full-dataset runs for values of K=1 through K=5. We repeated this subsetting procedure until all data subsets supported a value of K=1 using Pritchard and Wen's (2003) method for assessing the optimal value of K. We plotted consensus results from *CLUMPP* using the program *Distruct* (Rosenberg 2004).

We assessed patterns of isolation by distance (IBD; Wright 1943) within each study species using Mantel tests of geographic and genetic distance (Mantel 1967). Geographic distance between sampling sites was calculated as the linear shoreline distance between sites, using satellite images and GIS software (Manifold 2008). Populations in migration-drift equilibrium are expected to exhibit a positive linear relationship between genetic and geographic distance (Hutchison and Templeton 1999). The slope of the regression line describing this relationship is representative of the degree of gene flow between populations; as gene flow increases, the slope of the line will approach zero (Koizumi et al. 2006). The correlation between genetic and geographic distance will deteriorate as gene flow approaches zero and populations diverge by drift alone (Hutchison and Templeton 1999). Correlation between genetic and geographic distance will also be low in cases where there is a predominance of gene flow among populations leading to panmixia; in this case, all genetic distances would be close to zero. We plotted $F_{ST}/(1-F_{ST})$ versus geographic distance, as suggested by Rousset (1997) for one-dimensional (i.e. linear) habitats, to graphically assess patterns of IBD. We used Fischer's r-to-z transformation method to statistically compare correlations, as indicated by Cohen and Cohen (1983) and implemented using a web-based interface (Preacher 2002).

To assess the effects of individual loci on the observed IBD patterns, we also calculated the correlation between genetic and geographic distance for sub-sampled estimates of pairwise F_{ST} produced from jackknife analyses.

Although population structure can exist purely through IBD if dispersal is limited (Wright 1943), the signature of IBD can be confounded with the effect of landscape characteristics, as the number of potential barriers to gene flow will also increase as distance increases. Therefore, in order to examine the effects of landscape characteristics (in this case underwater substrate) on genetic divergence between populations, we used partial Mantel tests to account for geographic distance while testing for the effect of intervening sand substrate on genetic distance. Partial Mantel tests are closely related to partial regression, and seek to test the correlation between two distance matrices while controlling for the effect of a third matrix. We identified regions of near-shore sand substrate using field surveys and satellite imagery and mapped and measured these shoreline distances using GIS. Dissimilarity matrices were calculated as the total percentage of sand relative to rocky substrate for every pairwise distance between sites. Elevated type I error is a concern in using partial Mantel tests (Raufaste and Rousset 2001; Rousset 2002) because with spatially autocorrelated data, not all permutations of the data are equally likely. However, no straightforward analytical solutions yet exist to deal with this problem, and these tests seem to perform appropriately under most conditions (Castellano and Balletto 2002). Although results should be interpreted with caution, partial Mantel tests remain widely used in landscape genetic studies (Storfer et al. 2007). Simple and partial Mantel tests were performed using the package *ecodist* in R (Goslee and Urban 2007; R Development Core Team 2008), using 10,000 permutations to test for significance.

To compare patterns of population genetic structure among species, we used Mantel tests of pairwise φ'_{ST} matrices.

Results

Allelic variation and tests for equilibrium

We genotyped a total of 583 individuals from three study species at 12 microsatellite loci. For *S. diagramma*, all loci were polymorphic, for *P.* sp. "kazumbe" 10 of the 12 loci were polymorphic, and for *P.* sp. "moshi" 11 of the 12 loci were polymorphic (Table 1.1). Among polymorphic loci, number of alleles per locus ranges from 2 to 76 (Table 1.1), with an average across loci of 20, 14 and 31 for *P.* sp. "kazumbe"*, P.* sp. "moshi"*,* and *S. diagramma*, respectively.

Within populations of the three study species, there were no significant deviations from Hardy-Weinburg equilibrium at any locus. Across all populations within each species, there were significant deviations from Hardy-Weinberg equilibrium (*P.* sp. "kazumbe" $p = 0.010$; *P.* sp. "moshi" $p<0.001$; *S. diagramma* $p<0.001$), as would be expected from sampling across subpopulations (i.e. the Wahlund effect). Observed and expected heterozygosities per locus and population sample sizes are described in Table 1.2. For *S. diagramma*, average alleles per locus within populations ranged from 14-17, for *P.* sp. "kazumbe" they ranged from 8-11, and for *P.* sp. "moshi" they ranged from 4-7 (Table 1.2). There was no evidence for linkage disequilibrium among any pairs of loci in any population ($p > 0.05$).

Table 1.2. Observed and expected heterozygosities and average alleles per locus for *P.* sp. "kazumbe", *P.* sp. "moshi", and *S. diagramma* at each of the 8 study sites. N = sample size per population; Hexp = expected heterozygosity; Hnb = non-biased expected heterozygosity; Hobs = observed heterozygosity. No populations differ significantly from Hardy-Weinberg equilibrium expectations.

		P. sp. "kazumbe"					*P.* sp. "moshi"					*S.diagramma*				
Site	Name	N	Hexp	Hnb	Hobs	Ave alleles/loc	N	Hexp	Hnb	Hobs	Ave alleles/loc	N	Hexp	Hnb	Hobs	Ave alleles/loc
1	Gombe S	24	0.652	0.666	0.638	8.7	24	0.507	0.518	0.527	4.4	31	0.767	0.780	0.771	16.9
2	Katongwe N	31	0.645	0.656	0.642	9.9	0	--	--	--	--	24	0.785	0.801	0.826	14.7
3	Katongwe S	23	0.656	0.670	0.687	9.4	21	0.526	0.539	0.520	5.5	23	0.744	0.761	0.781	13.8
4	Kalalangabo	22	0.674	0.690	0.696	9.1	19	0.501	0.514	0.492	5.5	30	0.747	0.760	0.739	15.1
5	Nondwa	30	0.686	0.697	0.680	10.4	17	0.521	0.537	0.517	5.3	26	0.769	0.784	0.750	15.6
6	Hilltop	32	0.719	0.730	0.730	10.7	25	0.579	0.589	0.506	6.4	26	0.766	0.781	0.736	15.5
7	Jakobsen's	27	0.693	0.707	0.666	10.4	27	0.624	0.636	0.633	6.5	27	0.748	0.762	0.775	14.8
8	Ulambola	29	0.598	0.609	0.617	7.9	22	0.656	0.671	0.657	6.3	23	0.757	0.774	0.779	14.3

Population Structure

Both *Petrochromis* species were significantly differentiated among populations along the whole study shoreline (*P.* sp. "kazumbe" F_{ST} = 0.0727, p<0.001; *P.* sp. "moshi" F_{ST} = 0.157, p<0.001). For *S. diagramma* global F_{ST} also indicates significant differentiation, albeit with an F_{ST} value close to 0 (F_{ST} = 0.0057, p<0.001). Pairwise F_{ST}s, φ'$_{ST}$s, and maximum and minimum values of pairwise F_{ST} from the jackknife procedure over loci for each species are given in Appendix 1, Tables S2-S7. For *Petrochromis* species, a large majority of pairwise F_{ST} values were significant (23/28 for *P.* sp. "kazumbe"; 21/21 for *P.* sp. "moshi"), whereas for *Simochromis* 7 out of 28 pairwise comparisons produced F_{ST} values significantly different from zero after FDR correction for multiple tests. The largest F_{ST} for adjacent sites for both *Petrochromis* species was that between sites 7 and 8 (Appendix 1, Tables S1, S2). For *P.* sp. "kazumbe" none of the F_{ST} values between adjacent sites for sites 1-5 were significant. In contrast, all adjacent sites had significant F_{ST}s in *P.* sp. "moshi". In *S. diagramma,* two adjacent site pairs had significant F_{ST}s (4-5 and 5-6), but other site pairs had non-significant F_{ST}s, and all F_{ST} values were very close to 0 (Appendix 1, Table S4).

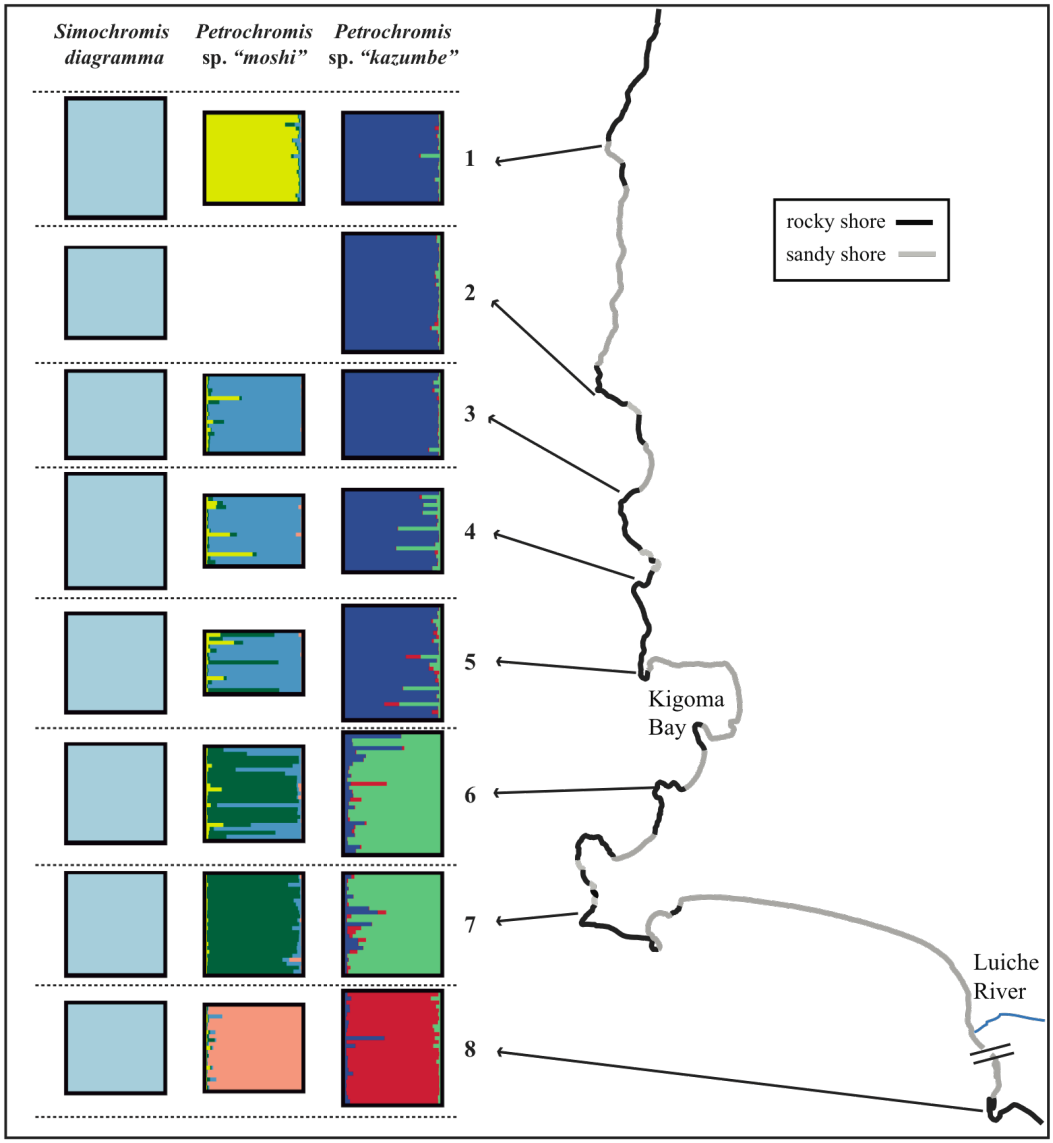


Figure 1.3. Bayesian analysis of population structure for three sympatric cichlid species. Each population of each species is represented by a colored box, wherein each horizontal linerepresents one individual in the data set and each color represents a different statistically-inferred genetic cluster (*K*). The x-axis describes the proportion of each individual's genotype (*Q*)belonging to each inferred genetic cluster. Plots shown are for the most likely values of *K* (number of genetic groups) for *S.* diagramma (far left, *K* = 1), *P.* sp. "moshi" (middle, K=4) andP. sp. "kazumbe" (far right, *K*=3).

STRUCTURE analyses support the patterns of differentiation indicated by F_{ST}s. For *P.* sp.

"kazumbe" the most likely value of *K* as assessed by lnP(D) values using Pritchard and Wen's

(2003) method was *K* equals 3 (Appendix 1, Figure S1). For *P.* sp. "moshi" the most likely *K* was

4 (Appendix 1, Figure S2). At these maximal values of *K*, deme boundaries are largely congruent between *Petrochromis* species, with genetic breaks present between sites 7 and 8 and between sites 5 and 6 for both species (Figure 1.3). For *P.* sp. "moshi" there is also a strong deme boundary between sites 1 and 3 (Figure 1.3). The lnP(D) results for *S. diagramma* supported $K = 1$ (Appendix 1, Figure S3). We obtained very similar results with hierarchical STRUCTURE analyses using Evanno et al.'s (2005) method (Appendix 1, Figure S4).

Isolation by distance and habitat influences on gene flow

There are significant positive relationships between genetic distance and geographic distance for *Petrochromis* species, but not for *S. diagramma*. Mantel tests of $F_{ST}/(1-F_{ST})$ and geographic distance indicate that there are highly significant isolation by distance patterns in both *Petrochromis* species (*P.* sp. "kazumbe", $p < 0.001$, $r = 0.943$, *P.* sp. "moshi", $p = 0.002$, $r = 0.905$) but not in *S. diagramma* ($p = 0.550$, $r = -0.238$) (Figure 1.4). These results are consistent across all sub-sampled estimates of pairwise F_{ST} from jackknife analyses (Figure 1.4; *P.* sp. "kazumbe", $p < 0.001$, $r > 0.940$ for all data subsets; *P.* sp. "moshi", $p < 0.005$, $r > 0.890$ for all data subsets; *S. diagramma* $p > 0.3$, $r = -0.174$ – -0.323 for data subsets), indicating that no single locus significantly alters the observed IBD patterns. There is no significant difference in the correlation of genetic and geographic distance between *Petrochromis* species ($p = 0.3908$), but between each *Petrochromis* species and *S. diagramma* there are significant differences in correlation ($p < 0.001$ for each comparison).

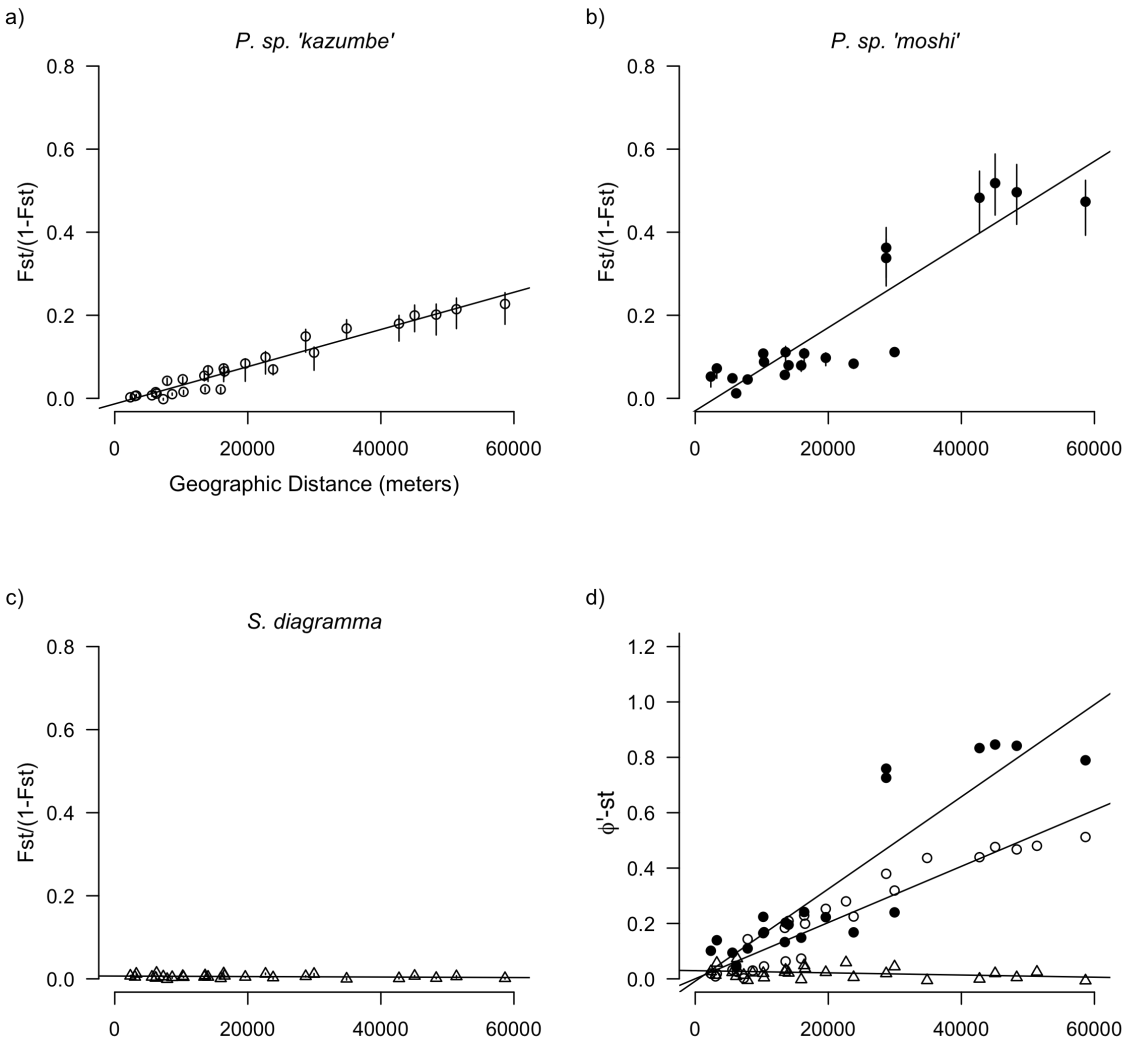


Figure 1.4. Isolation by distance patterns in *P.* sp. "kazumbe"*, P.* sp. "moshi" and *S. diagramma.*All plots show geographic distance (meters) on the x-axis, and genetic distance on the y-axis.Panels a-c: genetic distance as $F_{ST}/(1-F_{ST})$, vertical bars are maximum and minimum values derived from jackknife analyses of pairwise F_{ST} over loci; panel d: genetic distance as ϕ'_{ST}, a statistic describing population genetic structure and accounting for differences in effective population size and marker diversity among species. *P.* sp. "kazumbe" = open circle*, P.* sp."moshi" = closed circle, *S. diagramma* = triangle. *Simochromis diagramma* shows no evidence of geographic structure at this scale, and therefore no pattern of isolation by distance, whereas *Petrochromis* species show strong isolation by distance patterns.

Partial Mantel tests of F_{ST}, geographic distance, sand percentage between sites indicate that gene flow decreases in proportion to the percentage of sand between sites for *Petrochromis* species (*P.* sp. "kazumbe", $p < 0.001$, r = 0.961, *P.* sp. "moshi", $p = 0.002$, r = 0.849), but not for *S. diagramma* ($p = 0.938$, r = -0.030).

Mantel tests of pairwise φ'_{ST} matrices give support for significantly congruent population genetic structure between *Petrochromis* species ($p = 0.002$, r = 0.911), but not between *S. diagramma* and *P.* sp. "kazumbe" ($p = 0.507$, r = -0.258) nor between *S. diagramma* and *P.* sp. "moshi" ($p = 0.358$, r = -0.370). Using the standardized metric φ'_{ST} accounts for the substantially different levels of genetic diversity among species in making this comparison.

Discussion

We found significantly congruent patterns of fine-scale population structure between two species of the genus *Petrochromis* (Mantel tests of pairwise φ'_{ST} matrices, $p = 0.002$, r = 0.911) and no evidence for population structure in a third closely related species, *S. diagramma*. Both *Petrochromis* species showed strong spatial genetic structure within the study region (global F_{ST} = 0.073 and 0.157 for *P.* sp. "kazumbe" and *P.* sp. "moshi", respectively). Global F_{ST} estimates for these species are within the same orders of magnitude as the highest F_{ST}s documented in any East African cichlid species (Arnegard et al. 1999; Markert et al. 1999; Taylor et al. 2001; Rico et al. 2003). In contrast, *S. diagramma* exhibits no evidence of population genetic differentiation, even across substantial stretches of inappropriate habitat, suggesting large differences in levels of gene flow among populations for *Simochromis* and *Petrochromis* species. This is the first

demonstration that even sympatric, closely-related East African cichlids with shared habitat preferences can show dramatically different patterns of population structure at fine geographic scales.

Strong isolation by distance (IBD) patterns are present in both *Petrochromis* species. A positive and highly significant IBD pattern suggests that these species are in gene flow-drift equilibrium (Hutchison and Templeton 1999). Two factors may increase the variance observed in the relationship between genetic and geographic distance. First, the relative influence of drift should increase as distance between populations increases, creating greater variance in the relationship between genetic and geographic distance as geographic distance increases (Hutchison and Templeton 1999). Because the correlation between genetic and geographic distance is very high for both *Petrochromis* species ($r >= 0.9$), it is clear there is little increased variance due to the increased relative effects of drift at this spatial scale. Additionally, the presence of strong barriers to gene flow between some population pairs could increase the variance in the observed IBD relationship, because barriers would cause some populations to have diverged more than would be expected given the distance that separates them (Koizumi et al. 2006). Again, the very strong correlation between genetic and geographic distance suggests that barriers to gene flow do not obscure the strong IBD pattern for *Petrochromis* species. However, if there is a strong correlation between barriers to gene flow and geographic distance, it becomes difficult to tease apart the effects of distance alone from the accumulating effects of barriers to gene flow on observed genetic patterns. Among study sites in the Kigoma region there is a strong correlation between sand percentage between sites and geographic distance ($r=0.98$, $p<0.001$); it is possible that the

observed IBD patterns in *Petrochromis* species are, in fact, a result of the distribution of sand in this region rather than a pure IBD effect.

Partial mantel tests indicate that sand distance between sites explains the residual variance present in the genetic and geographic distance relationship for both *Petrochromis* species ($r = 0.961$ and $r = 0.849$ for *P.* sp. "kazumbe" and *P.* sp. "moshi", respectively). This is evidence that although gene flow in these species decreases with geographic distance alone, sand is a stronger inhibitor of gene flow than distance across rock substrate. Furthermore, the largest genetic discontinuities among adjacent population pairs in the *Petrochromis* species are explained by habitat features associated with long stretches of sandy shoreline. For both *Petrochromis* species, hierarchical STRUCTURE analyses and F_{ST} values indicate major differentiation between sites 7 and 8. This genetic break coincides with the Luiche River delta, a pervasive geological feature that has created a large expanse of sand extending several kilometers into the lake. River deltas are therefore formidable challenges to gene flow for *Petrochromis* species. Given that there are several rivers with larger deltaic systems that flow into the lake (e.g. Malagarasi, Ruzizi), this is a potentially important driver of population differentiation. A second barrier to gene flow suggested by STRUCTURE results for both *Petrochromis* species is Kigoma Bay, which falls between sites 5 and 6. Sandy bays of equivalent size to Kigoma Bay are common throughout the lake, and may be an important feature for the geographic subdivision of *Petrochromis* populations. Both of these findings suggest that expansive sand habitat functions as a substantial barrier to gene flow for *Petrochromis* species.

Petrochromis sp. "kazumbe" and *P.* sp. "moshi" STRUCTURE results are not congruent in one aspect: whereas STRUCTURE finds a deme boundary between populations 1 and 3 in *P.* sp.

"moshi", there is no evidence for population differentiation between populations 1, 2 or 3 in *P.* sp. "kazumbe" (Figure 1.3). Because insufficient numbers of individuals were present at site 2 to sample *P.* sp. "moshi" from this location, it is possible that the differentiation STRUCTURE finds between sites 1 and 3 is actually an IBD effect. Furthermore, if population sizes intervening sites 1 and 3 are very low for *P.* sp. "moshi", this may accentuate the IBD effect as it would decrease the effective number of migrants into the populations at sites 1 and 3, resulting in increased divergence between these populations.

The lack of genetic differentiation among populations of *S. diagramma* despite considerable stretches of sandy habitat is surprising. However, this fine-scale pattern is in concordance with large-scale phylogeographic analyses based on mitochondrial DNA for *S. diagramma*, in which there appeared to be no population genetic structure over the scale of hundreds of kilometers (Meyer et al. 1996). Non-significant IBD patterns can result from extremely low levels of gene flow between populations, in which populations diverge due to drift alone and pairwise genetic distances between populations are high, or from high levels of gene flow, in which case pairwise genetic distances would be close to 0 (Hutchison and Templeton 1999; Koizumi et al. 2006). Because pairwise F_{ST}s for *S. diagramma* are largely non-significant and close to 0 (Appendix 1, Table S4, Figure 1.4), the lack of IBD for this species suggests that there is extensive gene flow among *S. diagramma* populations, even across river deltas and sandy habitat that would be assumed inhospitable given their habitat preferences. The contrasting patterns of population genetic differentiation among *Simochromis* and *Petrochromis*, despite their shared habitat affinity and distribution, highlights the importance of considering intrinsic influences on population structure, such as ecological and behavioral traits, as well as extrinsic habitat features. Especially

in aquatic habitats, intrinsic traits can influence dramatically different patterns of population genetic structure among closely related species, even with shared extrinsic barriers to gene flow (Tringali et al. 1999; adult habitat preference: Rocha et al. 2002; life history characteristics: Whiteley et al. 2004; Zardoya et al. 2004).

Population structure and macroevolution

The idea that fine-scaled population genetic structure could lead to speciation is inherent in microallopatric explanations for the dramatic diversification of East African cichlids (e.g. Trewavas 1947; Fryer 1959; Fryer and Iles 1972). If population genetic structure is an important component of cichlid diversification, those groups with greater population genetic structure should be more diverse, while clades *without* fine-scale structure should be relatively depauperate. Our data are consistent with this hypothesized pattern: the genus *Petrochromis* is substantially more diverse than *Simochromis* based on information from widespread collections in the aquarium trade (*Petrochromis*: 18 recognized morphs, 6 described as species; *Simochromis*, 7 recognized morphs, 5 described as species; Brichard 1989; Herrmann 1996). However, a rigorous alpha taxonomy of both genera is needed. Even assuming that aquarium morphs are species or incipient species, our study provides just three data points in support of the link between population structure and diversity. Rigorous supporting evidence for fine-scale population structure as a driver of speciation in these cichlids will require additional information on three levels: population structure, rigorous alpha taxonomy and strong evidence for the monophyly of clades being compared.

No empirical work in any taxon has directly addressed the hypothesized link between population genetic subdivision and propensity to speciate (e.g. Vrba and Gould 1986; Rice 2004), although

some studies have examined the effects of traits potentially correlated with population genetic structure on species or subspecies richness (e.g. dispersal: Belliure et al. 2000; Phillimore et al. 2006; Smith and Farrell 2006). Information on scales of population genetic structure from multiple species in multiple clades, combined with robust assessments of species diversity in these groups, would provide a strong test for the link between diversity and population genetic structure. Of course, population genetic subdivision certainly does not always lead to speciation, and other factors likely influence the frequency at which subdivided populations speciate, such as degree of sexual selection (Barraclough et al. 1995; Mitra et al. 1996; Moller and Cuervo 1998; Owens et al. 1999; Arnqvist et al. 2000; Katzourakis et al. 2001) or capacity for phenotypic flexibility (Wimberger 1991; West-Eberhard 2003). In addition, even if taxa with fine-scale population structure are prone to speciation, it is also possible that these taxa would be prone to extinction, and that this genetic structure would produce no net increase in diversification rate (Coyne and Orr 2004). The best explanations for variation in clade diversity may thus involve multiple traits (e.g. Stuart-Fox and Owens 2003; Isaac et al. 2005; Phillimore et al. 2006).

Trophic specialization and diversification

The strikingly different patterns of population structure despite the sympatric distribution of *Petrochromis* and *Simochromis* species in this study area indicate that these species disperse in dramatically different ways within their shared habitat. One of the most prominent ecological differences between *Petrochromis* species and *S. diagramma* is the degree to which they are trophically specialized: although are both genera are generally reliant on algal grazing in rocky habitat, *S. diagramma's* ability to feed opportunistically may allow it to be more eurytopic than

Petrochromis, which could facilitate longer-distance dispersal. This is supported by frequent observations of *S. diagramma* over marginal and even sandy habitat, despite their prevalence within rocky habitat (C. Wagner personal observations; Brichard 1989; Meyer et al. 1996). Although the relationship between trophic ecology and population structure has not been examined explicitly in East African cichlids, the results of other studies focused on fine-scale population structure in cichlids also suggest its importance. For example, Danley et al. (2000) find little inhibition of gene flow across deep-water stretches in *Metriaclima zebra*, in contrast to previous studies of Malawi cichlids whose movement was strongly inhibited by deep water (Arnegard et al. 1999; Markert et al. 1999). The marked ecological difference between *Metriaclima* and the other studied species is that *Metriaclima* feeds in the water column whereas the other species have benthic feeding habits (Danley et al. 2000).

Because trophic ecology is a strong determinant of the way in which species use habitat, it likely has a substantial impact on population genetic structure. Links between specialization and population genetic structure have frequently been hypothesized, usually with the implication that such a link would promote speciation (Fryer 1959; Stanley 1979; Vrba 1984; Futuyma and Moreno 1988). Numerous studies have demonstrated differences in patterns of phylogeographic and population genetic structure based on differences in ecologically-based habitat preferences (e.g. Tringali et al. 1999; Rocha et al. 2002; Manier and Arnold 2006; Burridge et al. 2008), yet few focus on the impact of trophic specialization as it relates to habitat use and spatial genetic patterns. One exception is in studies of phytophagous insects, where host plant breadth provides a precise measure of ecological specialization with explicit spatial implications. However, no strong consensus has emerged from studies of the relationship between spatial genetic structure

and specialization in insects: Peterson and Denno (1998) find no evidence that trophic specialists exhibit stronger patterns of isolation by distance (IBD) than trophic generalists, although several studies involving fewer, more closely related insects do find such a correlation (Kelley et al. 2000; Smith and Farrell 2006). Future work should assess the relationship between trophic specialization and population genetic structure in other taxa, where this topic remains largely unstudied, with explicit effort to address the hypothesized links to patterns of species diversity.

Trophic specialization has figured prominently as an explanatory factor in cichlid diversification through the much cited idea that jaw flexibility was a "key innovation" (Liem 1973) leading to explosive diversification of African cichlids. While this idea conveys the capacity for functional diversification of cichlid jaws, it does not explicitly link trophic flexibility to the speciation process. Trophic versatility does not necessarily promote divergence (Danley and Kocher 2001) or the evolution of reproductive isolation between trophic morphs. However, if trophic ecology affects population genetic structure, and if such population structure influences the probability of speciation, then trophic specialization itself could be causally related to speciation. In this way, trophic specialization, a major feature of the renowned adaptive nature of cichlid radiations, could have facilitated diversification in East African cichlids.

Acknowledgements

This chapter was written in collaboration with Amy R. McCune. Many thanks to Rick Harrison, Irby Lovette, Ellinor Michel, Dan Rabosky and the Lovette lab group for comments on this manuscript, to Pete McIntyre and Jon Todd for helpful discussions, to Laura Stenzler and Chris Makarewich for technical consultation during lab work, to Aurelie Coulon for analytical advice, and to Keith Jenkins for help with GIS work. George Kazumbe provided essential help with fish collection, and the NSF Nyanza Project, the Tanzanian Fisheries Research Institute, the Cornell University Museum of Vertebrates and the Fuller Evolutionary Biology Program at the Cornell Lab of Ornithology provided logistical support. We thank Michael Hellberg and two anonymous reviewers for helpful comments on previous versions of the manuscript. Funding for this study was provided by SSB Research Award, ASIH Raney Fund, ACA Jordan Endowment Fund, Sigma Xi, Cornell University Department of EEB and Cornell University Mellon Fund grants to CEW, and NSF DDEP grant INT-0724247 to CEW and ARM. In addition, NSF grants ATM-0223920 and DBI-0353765 (The Nyanza Project) provided logistical support in the field. Part of this work was carried out using the resources of the Computational Biology Service Unit at Cornell University, which is partially funded by Microsoft Corporation.

REFERENCES

Ackermann, M., and M. Doebeli. 2004. Evolution of niche width and adaptive diversification. Evolution 58:2599-2612.

Adams, C. E., D. J. Hamilton, I. McCarthy, A. J. Wilson, A. Grant, G. Alexander, S. Waldron, S. S. Snorasson, M. M. Ferguson, and S. Skulason. 2006. Does breeding site fidelity drive phenotypic and genetic sub-structuring of a population of Arctic charr? Evol. Ecol. 20:11-26.

Arnegard, M. E., and A. S. Kondrashov. 2004. Sympatric speciation by sexual selection alone is unlikely. Evolution 58:222-237.

Arnegard, M. E., J. A. Markert, P. D. Danley, J. R. Stauffer, A. J. Ambali, and T. D. Kocher. 1999. Population structure and colour variation of the cichlid fish *Labeotropheus fulleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. Proc. R. Soc. B 266:119-130.

Arnqvist, G., M. Edvardsson, U. Friberg, and T. Nilsson. 2000. Sexual conflict promotes speciation in insects. Proc. Natl. Acad. Sci. USA 97:10460-10464.

Baric, S., W. Salzburger, and C. Sturmbauer. 2003. Phylogeography and evolution of the Tanganyikan cichlid genus Tropheus based upon mitochondrial DNA sequences. J. Mol. Evol. 56:54-68.

Barluenga, M., and A. Meyer. 2004. The Midas cichlid species complex: incipient sympatric speciation in Nicaraguan cichlid fishes? Mol. Ecol. 13:2061-2076.

Barluenga, M., and A. Meyer. 2005. Old fish in a young lake: stone loach (Pisces: *Barbatula barbatula*) populations in Lake Constance are genetically isolated by distance. Mol. Ecol. 14:1229-1239.

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

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 1996-2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Belliure, J., G. Sorci, A. P. Moller, and J. Clobert. 2000. Dispersal distances predict subspecies richness in birds. J. Evol. Biol. 13:480-487.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B 57:289-300.

Bergek, S., and M. Bjorklund. 2007. Cryptic barriers to dispersal within a lake allow genetic differentiation of Eurasian perch. Evolution 61:2035-2041.

Bolnick, D. I. 2004. Waiting for sympatric speciation. Evolution 58:895-899.

Brichard, P. 1989. Pierre Brichard's book of cichlids and all the other fish of Lake Tanganyika. T.F.H. Publications, New Jersey.

Brownstein, M. J., J. D. Carpten, and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. . Bio Techniques 20:1004-1010.

Burger, R., and K. A. Schneider. 2006. Intraspecific competitive divergence and convergence under assortative mating. Am. Nat. 167:190-205.

Burridge, C. P., D. Craw, D. C. Jack, T. M. King, and J. M. Waters. 2008. Does fish ecology predict dispersal across a river drainage divide? Evolution 62:1484-1499.

Castellano, S., and E. Balletto. 2002. Is the partial mantel test inadequate? Evolution 56:1871-1873.

Castric, V., and L. Bernatchez. 2004. Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase in genetic differences with distance. Mol. Ecol. 13:1299-1312.

Cohen, A. S., M. Soreghan, and C. Scholz. 1993. Estimating the age of ancient lake basins: An example from Lake Tanganyika. Geology 21:511-514.

Cohen, A. S., J. R. Stone, K. R. M. Beuning, L. E. Park, P. N. Reinthal, D. Dettman, C. Scholz, T. C. Johnson, J. W. King, M. R. Talbot, E. T. Brown, and S. J. Ivory. 2007. Ecological consequences of early Lake Pleistocene megadroughts in tropical Africa. Proc. Natl. Acad. Sci. USA 104:16422-16427.

Cohen, J., and P. Cohen. 1983. Applied multiple regression/correlation analysis for the behavioral sciences. Erlbaum, Hillsdale, NJ.

Coulon, A., J. W. Fitzpatrick, R. Bowman, B. M. Stith, A. Makarewich, L. M. Stenzler, and I. J. Lovette. 2008. Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). Mol. Ecol. 17:1685-1701.

Coulter, G. W. 1991. Lake Tanganyika and its Life. Oxford University Press, London.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Danley, P. D., and T. D. Kocher. 2001. Speciation in rapidly diverging systems: lessons from Lake Malawi. Mol. Ecol. 10:1075-1086.

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

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

Egger, B., S. Koblmuller, C. Sturmbauer, and K. M. Sefc. 2007. Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. BMC Evol. Biol. 7:137.

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

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

Fryer, G. 1959. Some aspects of evolution in Lake Nyasa. Evolution 13:440-451.

Fryer, G., and T. D. Iles. 1972. The cichlid fishes of the great lakes of Africa: their biology and evolution. Tropical Fish Hobbyist Publications Inc., Edinburgh.

Futuyma, D. J., and G. Moreno. 1988. The evolution of ecological specialization. Annu. Rev. Ecol. Syst. 19:207-233.

Gavrilets, S. 2003. Models of speciation: what have we learned in 40 years? Evolution 57:2197-2215.

Goslee, S. C., and D. L. Urban. 2007. The ecodist package for dissimilarity-based analysis of ecological data. Journal of Statistical Software 22:1-19.

Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes 5:184-186.

Goudet, J. 2006. HIERFSTAT: Estimation and tests of hierarchical F-statistics. R package version 0.04-4.

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hailer, F., B. Gautschi, and B. Helander. 2005. Development and multiplex PCR amplification of novel microsatellite markers in the White-tailed Sea Eagle, *Haliaeetus albicilla* (Aves: Falconiformes, Accipitridae). Mol. Ecol. Notes 5:938-940.

Haldane, J. B. S. 1954. An exact test for randomness of mating. Journal of Genetics 52:631-635.

Hedrick, P. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.

Herrmann, H. 1996. Cichlids from Lake Tanganyika. Dahne Verlag, Ettlingen.

Hutchison, D. W., and A. R. Templeton. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. Evolution 53:1898-1914.

Isaac, N. J. B., K. E. Jones, J. L. Gittleman, and A. Purvis. 2005. Correlates of species richness in mammals: body size, life history, and ecology. . Am. Nat. 165:600-607.

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

Katzourakis, A., A. Purvis, S. Azmeh, G. Rotheray, and F. Gilbert. 2001. Macroevolution of hoverflies (Diptera: Syrphidae): the effect of using higher-level taxa in studies of biodiversity, and correlates of species richness. J. Evol. Biol. 14:219-227.

Kelley, S. T., B. D. Farrell, and J. B. Mitton. 2000. Effects of specialization on genetic differentiation in sister species of bark beetles. Heredity 84:218-227.

Kimura, M., and G. H. Weiss. 1964. The stepping-stone model of population structure and the decrease of genetic correlation with distance. . Genetics 49:561-576.

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

Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nature Reviews Genetics 5:288-298.

Koizumi, I., S. Yamamoto, and K. Maekawa. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden charr. Mol. Ecol. 15:3175-3189.

Liem, K. F. 1973. Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Systematic Zoology 22:425-441.

Manier, M., and S. Arnold. 2006. Ecological correlates of population genetic structure: a comparative approach using a vertebrate metacommunity. Proc. R. Soc. B 273:3001-3009.

Manifold. 2008. Manifold System 8.0. CDA International Ltd.

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

Markert, J. A., M. E. Arnegard, P. D. Danley, and T. D. Kocher. 1999. Biogeography and population genetics of the Lake Malawi cichlid Melanochromis auratus: habitat transience, philopatry and speciation. Mol. Ecol. 8:1013-1026.

Meirmans, P. G. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution 60:2399-2402.

Meirmans, P. G., and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol. Ecol. Notes 4:792-794.

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

Mitra, S., H. Landel, and S. Pruett Jones. 1996. Species richness covaries with mating system in birds. Auk 113:544-551.

Moller, A. P., and J. J. Cuervo. 1998. Speciation and feather ornamentation in birds. Evolution 52:859-869.

Owens, I. P. F., P. M. Bennett, and P. H. Harvey. 1999. Species richness among birds: body size, life history, sexual selection or ecology? Proc. R. Soc. B 266:933-939.

Peterson, M. A., and R. F. Denno. 1998. The influence of dispersal and diet breadth on patterns of genetic isolation by distance in phytophagous insects. Am. Nat. 152:428-446.

Phillimore, A., R. Freckleton, C. D. Orme, and I. F. Owens. 2006. Ecology Predicts Large-Scale Patterns of Phylogenetic Diversification in Birds. Am. Nat. 168:220-229.

Pounds, S. B. 2006. Estimation and control of multiple testing error rates for microarray studies. Briefings in Bioinformatics 7:25-36.

Preacher, K. J. 2002. Calculation for the tests of the difference between two independent correlation coefficients. Available from http://www.quantpsy.org.

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

Pritchard, J. K., and W. Wen. 2003. Documentation for STRUCTURE software: Version 2. . Available from http://pritch.bsd.uchicago.edu.

Raufaste, N., and F. Rousset. 2001. Are partial mantel tests adequate? Evolution 55:1703-1705.

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

Rice, S. H. 2004. Evolutionary theory: mathematical and conceptual foundations. Sunderland, Sunderland, MA.

Rico, C., P. Bouteillon, M. J. H. Van Oppen, M. E. Knight, G. M. Hewitt, and G. F. Turner. 2003. No evidence for parallel sympatric speciation in cichlid species of the genus *Pseudotropheus* from north-western Lake Malawi. J. Evol. Biol. 16:37-46.

Rocha, L. A., A. L. Bass, D. R. Robertson, and B. W. Bowen. 2002. Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). Mol. Ecol. 11:243-252.

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

Rousset, F. 1997. Genetic differentiation and estimation of gene flow using F-statistics under isolation by distance. Genetics 145:1219-1228.

Rousset, F. 2002. Partial mantel tests: reply to Castellano and Balletto. Evolution 56:187-1875.

Rundle, H., and P. Nosil. 2005. Ecological Speciation. Ecol. Lett. 8:336-352.

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

Schliewen, U. K., D. Tautz, and S. Paabo. 1994. Sympatric Speciation Suggested By Monophyly Of Crater Lake Cichlids. Nature 368:629-632.

Schluter, D. 1996. Ecological speciation in postglacial fishes. Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences 351:807-814.

Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

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

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

Slatkin, M. 1993. Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47:264-279.

Smith, C. I., and B. D. Farrell. 2006. Evolutionary consequences of dispersal ability in cactus-feeding insects. Genetica 126:323-334.

Soreghan, M., and A. S. Cohen. 1996. Textural and compositional variability across littoral segments of Lake Tanganyika: the effect of asymmetric basin structure on sedimentation in large rift lakes. AAPG Bulletin 80:382-409.

Stanley, S. M. 1979. Macroevolution: pattern and process. W.H. Freeman, San Francisco.

Storfer, A., M. A. Murphy, J. S. Evans, C. S. Goldberg, S. Robinson, S. F. Spear, R. Dezzani, E. Delmelle, L. Vierling, and L. P. Waits. 2007. Putting the 'landscape' in landscape genetics. Heredity 98:128-142.

Stuart-Fox, D., and I. P. F. Owens. 2003. Species richness in agamid lizards: chance, body size, sexual selection or ecology? J. Evol. Biol. 16:659-669.

Sturmbauer, C., S. Koblmuller, K. M. Sefc, and N. Duftner. 2005. Phylogeographic history of the genus Tropheus, a lineage of rock-dwelling cichlid fishes endemic to Lake Tanganyika. Hydrobiologia 542:335-366.

Sturmbauer, C., and A. Meyer. 1992. Genetic-Divergence, Speciation And Morphological Stasis In A Lineage Of African Cichlid Fishes. Nature 358:578-581.

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

Team, R. D. C. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Tiercelin, J. J., and A. Mondeguer. 1991. The geology of the Tanganyika trough. Pp. 7-48 *in* G. W. Coulter, ed. Lake Tanganyika and its Life. Oxford University Press, London.

Trewavas, E. 1947. Speciation in cichlid fishes of East African Lakes. Nature 160:96-97.

Tringali, M. D., T. M. Bert, S. Seyoum, E. Bermingham, and D. Bartolacci. 1999. Molecular phylogenetics and ecological diversification of the transisthmian fish genus *Centropomus* (Perciformes: Centropomidae). Mol. Phylogen. Evol. 13:193-207.

Van Oppen, M. J. H., G. F. Turner, C. Rico, J. C. Deutsch, K. M. Ibrahim, R. L. Robinson, and G. M. Hewitt. 1997. Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. Proc. R. Soc. B 264.

Verheyen, E., L. Ruber, J. Snoeks, and A. Meyer. 1996. Mitochondrial phylogeography of rock-dwelling cichlid fishes reveals evolutionary influence of historical lake level fluctuations of Lake Tanganyika, Africa. Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences 351:797-805.

Vrba, E. S. 1984. Evolutionary pattern and process in the sister-group Alcelaphini-Aepycerotini (Mammalia: Bovidae). Pp. 247-271 *in* A. K. Behrensmeyer, and A. P. Hill, eds. Fossils in the Making. University of Chicago Press, Chicago.

Vrba, E. S., and S. J. Gould. 1986. The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. Paleobiology 12:217-228.

Weir, B. S. 1996. Genetic Data Analysis II. Sinauer, Sunderland, Mass.

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

West-Eberhard, M. J. 2003. Developmental plasticity and evolution. Oxford University Press, Oxford.

Whiteley, A. R., P. Spruell, and F. W. Allendorf. 2004. Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. Mol. Ecol. 13:3675-3688.

Wiens, J. J. 2004. Speciation and ecology revisited: phylogenetic niche conservatism and the origin of species. Evolution 58:193-197.

Wimberger, P. H. 1991. Plasticity of jaw and skull morphology in the neotropical cichlids *Geophagus brasilliensis* and *G. Stradochneri*. Evolution 45:1545-1563.

Wright, S. 1943. Isolation by distance. Genetics 28:114-138.

Yamaoka, K. 1983. Feeding behaviour and dental morphology of algae scraping cichlids (Pisces: Teleostei) in Lake Tanganyika. African Study Monographs 4:77-89.

Zardoya, R., R. Castilho, C. Grande, L. Favre-Krey, S. Caetano, S. Marcato, G. Krey, and T. Patarnello. 2004. Differential population structuring of two closely realted fish species, the mackerel (*Scomber scombrus*) and the chub mackerel (*Scomber japonicus*), in the Mediterranean Sea. Mol. Ecol. 13:1785-1798.

CHAPTER 2

RECENT SPECIATION IN SYMPATRIC TANGANYIKAN CICHLID COLOR MORPHS

Abstract

Lake Tanganyika, Africa's oldest lake, harbors an impressive diversity of cichlid fishes. Although diversification in its radiating groups is thought to have been initially rapid, cichlids from Lake Tanganyika show little evidence for ongoing speciation. In contrast, examples of recent divergence among sympatric color morphs are well known in haplochromine cichlids from Lakes Malawi and Victoria. Here we report genetic evidence for recent divergence between two sympatric Tanganyikan cichlid color morphs. These *Petrochromis* morphs share mitochondrial haplotypes, yet microsatellite loci reveal that their sympatric populations form distinct genetic groups. Nuclear divergence between the two morphs is equivalent to that which arises geographically within one of the morphs over short distances, and is substantially smaller than that among other sympatric species in this genus. These patterns suggest that these morphs diverged only recently, yet that barriers to gene flow exist which prevent extensive admixture despite their sympatric distribution. The morphs studied here provide an unusual example of active diversification in Lake Tanganyika's generally ancient cichlid fauna and enable comparisons of speciation processes between Lake Tanganyika and other African lakes.

Introduction

The cichlid fish radiations of East Africa are hailed as the most diverse vertebrate species flocks (Kocher 2004), and as the animal group with the fastest-known speciation rates (McCune 1997). Less appreciated, however, are the dramatic differences among cichlid radiations in different African lakes, among which Lake Tanganyika is an outlier owing to its extremely old age (9-12 million years; Cohen et al. 1993). In contrast with the single-lineage origins of cichlid diversity in Lakes Malawi and Victoria (but see Seehausen et al. 2003; Joyce et al. 2010), the Tanganyikan cichlid fauna descends from eight colonizing lineages (Salzburger et al. 2002; Salzburger et al. 2005). Species richness in Lake Tanganyika is substantially lower than in Lakes Malawi and Victoria (Tanganyika: ~250 spp.; Malawi: 451-600 spp.; Victoria: 447-535 spp.; Genner et al. 2004), and diversification rates in Tanganyika are much lower than in the recent, rapid radiations of Malawi and Victoria (Day et al. 2008). Lake Tanganyika is considered the ancestral source of the riverine lineages that seeded the haplochromine radiations of both Malawi and Victoria (Salzburger et al. 2005). However, unlike these radiations, where sharing of polymorphism among species is rampant due to extremely recent divergence and ongoing hybridization (e.g. Meyer et al. 1990; Nagl et al. 1998; Won et al. 2006), Lake Tanganyika's cichlid groups are older clades in which no cases of incomplete lineage sorting of mitochondrial DNA at the species level have been described (Sturmbauer et al. 2003).

The existence of sympatric, closely-related color morphs in Lakes Malawi and Victoria has spurred interest in these fishes as potential cases of sympatric speciation (Seehausen and van Alphen 1999). In contrast, although geographic differences in color are common in Tanganyikan

cichlids (Kohda et al. 1996) and are well-studied in the genus *Tropheus* (Baric et al. 2003; Egger et al. 2007), no sympatric color morphs have been described from Lake Tanganyika (but for a case of artificial sympatry see Salzburger et al. 2006). Tanganyika's cichlids thus represent an important contrast to the haplochromines of Lakes Malawi and Victoria. Do cichlid speciation mechanisms differ in Lake Tanganyika compared to the younger radiations of Lakes Malawi and Victoria? Or do the differences in diversity arise from speciation-independent mechanisms, like the winnowing of ecologically similar forms through extinction after speciation? Understanding how processes of speciation in Tanganyika's cichlids compare and contrast with those of other lakes will provide insight into the interplay between diversity generating mechanisms and long-term ecological processes in determining patterns of extant diversity in African cichlids.

The crucial questions for understanding speciation in sexually reproducing organisms, under any geographic scenario, are 1) what maintains reproductive isolation between recently separated species, and 2) how did these isolating mechanisms originate? Although full tests of these questions in wild populations are exceedingly difficult, our best approaches involve careful study of very recently diverged species and of populations in the process of divergence (Coyne and Orr 2004; Maan et al. 2004; Seehausen et al. 2008). Here we investigate genetic relationships between sympatrically distributed color morphs of the genus *Petrochromis* in the region of Kigoma, Tanzania.

Study System

Petrochromis cichlids are members of the tribe Tropheini, the Tanganyikan lineage that is sister to the clade that includes the haplochromine radiations of Lakes Malawi and Victoria. Like cichlids from the radiations in Malawi and Victoria, tropheines maternally mouthbrood their

young and are generally polygamous (Brichard 1989). However, other characteristics of the tropheine radiation contrast substantially with the Malawi and Victoria haplochromine radiations. Tropheine cichlids are not nearly as diverse (24 spp; Koblmuller et al. 2010), nor do they show the dramatic sexual dimorphism and color polymorphism common in haplochromines from Lakes Malawi and Victoria (Seehausen et al. 1999a; Dijkstra et al. 2009; Roberts et al. 2009). In previous work, we demonstrated that *P.* sp. "kazumbe" and sympatric *P.* sp. "moshi", both well-recognized morphs in the aquarium trade, have strongly and congruently geographically structured populations in the Kigoma region of Lake Tanganyika (Wagner and McCune 2009). Specifically, we showed that genetic breaks in *P.* sp. "kazumbe" populations coincide with two long stretches of sandy habitat, the Kigoma Bay and the Luiche River delta, subdividing this species into northern, mid, and southern genetic groups within the region (Wagner and McCune 2009). Another *Petrochromis* morph, *Petrochromis* cf. *polyodon* (Boulenger 1898), occurs in sympatry with *P.* sp. "kazumbe" in the Kigoma region. Adults of *P.* sp. "kazumbe" and *P.* cf. *polyodon* look very similar, but are readily distinguishable based on the amount of orange coloration on the body, head and fins: *P.* sp. "kazumbe" has a mostly orange coloration with a light gray background whereas *P.* cf. *polyodon* is predominantly light gray-blue with a limited amount of orange coloration. Given the sympatric distributions of *P.* cf. *polyodon* and *P.* sp. "kazumbe," and the importance of coloration in other cichlid groups as a cue for assortative mating, we here investigate whether these color morphs represent polymorphism within panmictic populations, or whether there is population genetic evidence for assortative mating between these sympatric morphs. Using mitochondrial DNA sequence data and multi-locus microsatellite genotypes, we tested for evidence of genetic differentiation between these groups

via haplotype analyses and Bayesian analyses of population structure. To provide a comparative scale for the small magnitude of genetic differentiation of these two morphs, we conducted phylogenetic analyses based on sequence data collected from all tropheine cichlids present in the Kigoma region and population genetic analyses based on microsatellite data from their close *Petrochromis* relatives.

Materials and Methods

Sample collection

We collected *P.* sp. "kazumbe" (hereafter the "orange morph") at eight sites spanning 60 kilometers of coastline in the Kigoma region of Tanzania in 2005 and 2007 (Appendix 2, Table S1). We also collected 11 individuals that we identified as the as *Petrochromis* cf. *polyodon* (hereafter the "blue morph") at four of these sites on the same visits. We collected no juveniles (<10 cm standard length) with the distinctive blue morph-type coloration; all juveniles collected had orange morph-type coloration. Between 1 and 10 meters depth, orange morph *Petrochromis* are far more common than are blue morph *Petrochromis* (C. Wagner, unpublished data), although their relative abundance at greater depths is unknown. As a metric of genetic divergence among other sympatric close relatives, we collected samples of the four other *Petrochromis* species that are found sympatrically with the orange and blue morph *Petrochromis*. These included *P. orthognathus, P. famula,* and the undescribed species *P.* sp. "moshi" and *P.* sp. "green". Additionally, representatives from all other tropheine cichlid species found in the Kigoma region (nine species in seven genera) were collected at a subset of these sites in 2002, 2005 and 2007. All fishes were collected using gill nets while snorkeling in the rocky littoral zone, from

approximately 1-10 meters depth. Fin clips were preserved in DMSO-EDTA buffer (Seutin and White 1991) for genetic work. All specimens were retained as vouchers and have been deposited in the Cornell Museum of Vertebrates (Appendix 2, Table S1).

DNA extraction, PCR, Sequencing, Genotyping

We extracted genomic DNA from fin clips using DNeasy Tissue Kits (Qiagen). We PCR-amplified and sequenced the mitochondrial nitrogen dehydrogenase subunit 2 gene (ND2, 1,047 bp) for 19 individuals of the orange morph and 17 individuals of the blue morph, including juveniles, using the primers and conditions described in Wagner et al. (2009). We also sequenced ND2 and the mitochondrial cytochrome B oxidase gene (cytB, 1,149 bp) using methods described in Wagner et al. (2009), from all other tropheine cichlid species and morphs found in the Kigoma region.

We amplified 12 microsatellite loci previously developed for Tanganyikan and Malawiian cichlid species (Appendix 2, Table S2) for orange and blue morph *Petrochromis*. Using the same loci, we genotyped individuals of the other sympatric undescribed *Petrochromis* species: *P.* sp. "moshi" (see also Wagner and McCune 2009) and *P.* sp. "green". Details of the PCR and genotyping reactions are described in Wagner & McCune (2009). We scored microsatellite genotypes using Genemapper version 3.7 (Applied Biosystems) and verified each genotype by eye. For the orange and blue morph *Petrochromis*, all samples were genotyped and scored blind to the field species identification of the sample, and all scoring used the same panels and bin sets for all loci and all individuals.

Data analysis

To place the relationship of the orange and blue *Petrochromis* morphs in a broader phylogenetic context, we conducted a phylogenetic analysis using Mr. Bayes 3.1.2 (Huelsenbeck et al. 2001; Ronquist and Huelsenbeck 2003) of ND2 and cytB sequences representing the full Kigoma region tropheine cichlid community. We ran Mr. Bayes for two runs of 10 million generations each, and discarded the first 10% of trees to account for burn-in, leaving the posterior distribution of trees with standard deviations of split frequencies below 0.01. We used TCS (Clement et al. 2000) analyses of additional ND2 sequences from the orange and blue morphs to explore the patterns of mitochondrial variation among these morphs.

Using microsatellite genotypes, we employed the Bayesian clustering program STRUCTURE v. 3.3 (Pritchard et al. 2000) to find the most likely number of genetic groups (*K*) for the complete set of orange and blue morph individuals. We ran STRUCTURE for values of *K* from 1 through 10, for 10 runs at each *K* value. Each run consisted of 180,000 generations for burn-in and 1 million generations post-burn-in, used the admixture and correlated allele frequency models, and were set to infer the parameter alpha separately for each population. We then assessed the most likely number of genetic groups by examining the value of *K* for which LnP(D) values reached their maximum (Pritchard and Wen 2003). We used the program CLUMPP (Jakobsson and Rosenberg 2007) to summarize results over multiple runs and *Distruct* (Rosenberg 2004) to plot STRUCTURE results for the most likely *K* value.

Using microsatellite data, we calculated F_{ST} values for all pairs of population samples identified through STRUCTURE analyses. We also calculated pairwise differences between *P.* sp. “moshi”, *P.* sp. “green” and the orange and blue morph groups. We used Weir and Cockerham’s (1984)

estimator and tested for significance using permutation tests of 10,000 replicates in GENETIX 4.05.2 (Belkhir et al. 1996-2004). To evaluate the effects of individual loci on estimates of F_{ST} we performed jackknife analyses over loci, as described in Wagner and McCune (2009).

Results

Phylogenetic analyses of mtDNA sequences for the tropheine cichlid communities sampled in this study support the grouping of four *Petrochromis* morphs found in the Kigoma region: *P.* sp. "green", *P.* sp. "moshi", and the orange and blue *Petrochromis* morphs (Figure 2.1a). The orange and blue morphs cannot be distinguished based on mtDNA sequence data. As expected of well differentiated species, *P.* sp. "green" and *P.* sp. "moshi" individuals each form strongly supported monophyletic groups, whereas the orange and blue morph *Petrochromis* together form a strongly supported but intermixed haplotype clade (Figure 2.1). This haplotype mixing contrasts with the 11 additional tropheine species in the analysis, all of which are represented by multiple sequences and show strong evidence for reciprocal monophyly with respect to all other species (Appendix 2, Figure S1). Phylogenetic resolution of basal relationships in the tree is generally poor (Appendix 2, Figure S1), as has been shown in other phylogenetic studies of tropheine cichlids and interpreted as an indication of ancient rapid radiation and hybridization (Sturmbauer et al. 2003; Koblmuller et al. 2010).

Eleven ND2 haplotypes were recovered from 36 orange and blue morph individuals. Twenty-two of these individuals shared one common haplotype, three other haplotypes were shared between morphs, and the only haplotypes not shared among morphs were haplotypes recovered from

single individuals (Figure 2.1b). All sequences have been deposited in Genbank (Appendix 2, Table S1).

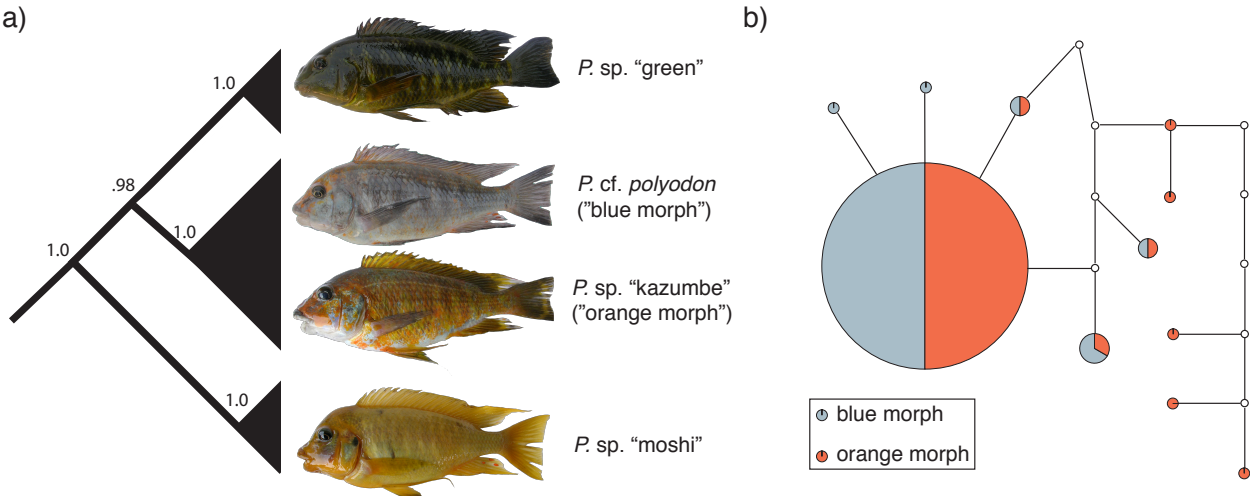


Figure 2.1. a) Two undescribed sympatric *Petrochromis* species plus the orange and blue *Petrochromis* morphs form a strongly supported mitochondrial clade. Whereas *P.* sp. "green" and *P*. sp. "moshi" are reciprocally monophyletic, the orange and blue *Petrochromis* morphs share mitochondrial haplotypes. b) Relationships among ND2 haplotypes recovered from orange and blue morph *Petrochromis* individuals. Circles are proportional to the number of individuals with a given haplotype.

Eleven of the 12 microsatellite loci were variable for the orange and blue *Petrochromis* morphs and were used in further analyses. STRUCTURE runs supported four genetic groups among orange and blue morph genotypes (Figure 2.2; Appendix 2, Figure S2). Three of these represent spatial genetic structure from the extensively sampled orange morph populations, forming "north", "mid" and "south" genetic groups, as described in Wagner & McCune (2009). The fourth genetic group includes all fish that were identified as the blue morph in the field, all of which were adults greater than 10 cm in standard length (n = 11). All of these fish had assignment probabilities greater than 95% to this fourth genetic group. This fourth genetic group additionally includes juvenile fish not recognized as the blue morph in the field (n = 13). Three

juvenile fish in total assigned to the fourth genetic group had assignment probabilities less than 88% (60%, 62% and 72%). One fish was split 38%-33%-28% to the first, second and fourth genetic groups, respectively.

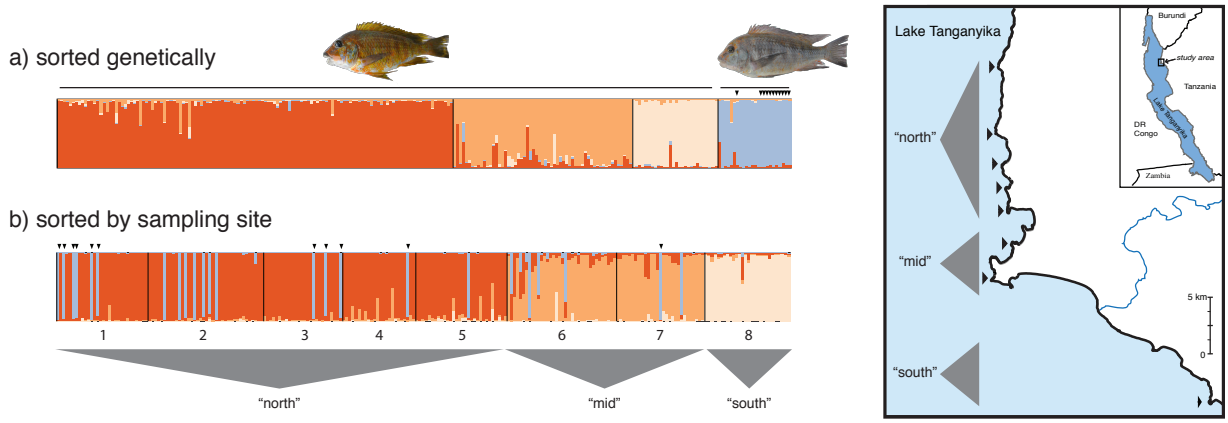


Figure 2.2. STRUCTURE analyses support the existence of four genetic groups (*K*=4) among individuals from eight sampling sites. Three of these groups are geographic divergence within the orange morph; the fourth is blue morph individuals in sympatry with orange morph individuals. This demonstrates that orange morph and blue morph populations are not panmictic, despite their sympatric distribution. In a) individuals are sorted by genetic group; in b) individuals are sorted by sampling site (1-8). Within the orange morph there are three geographically distributed groups, "north", "mid", and "south" (corresponding to the three shades of orange), as reported in Wagner and McCune (2009). Blue morph individuals (blue), regardless of sampling location, form the fourth genetic group. Black triangles correspond to individuals identified in the field as the blue morph. Individual assignment probabilities shown here are the consensus generated by CLUMPP from 10 STRUCTURE runs of *K*=4.

All pairwise F_{ST}s between STRUCTURE-assigned populations were highly significant. Pairwise F_{ST} values were highest in comparisons of the blue morph to the orange morph geographic subset populations (Appendix 2, Figure S3). The pairwise F_{ST} value for the northernmost versus southernmost orange morph populations (Wagner and McCune 2009) is nearly equivalent to the pairwise F_{ST} value for pooled orange morph geographic populations versus the blue morph (0.19

vs. 0.21; Figure 2.3). All other pairwise F_{ST} values between morphs are substantially higher (> 0.3) than the value for orange morph versus blue morph individuals (Figure 2.3).

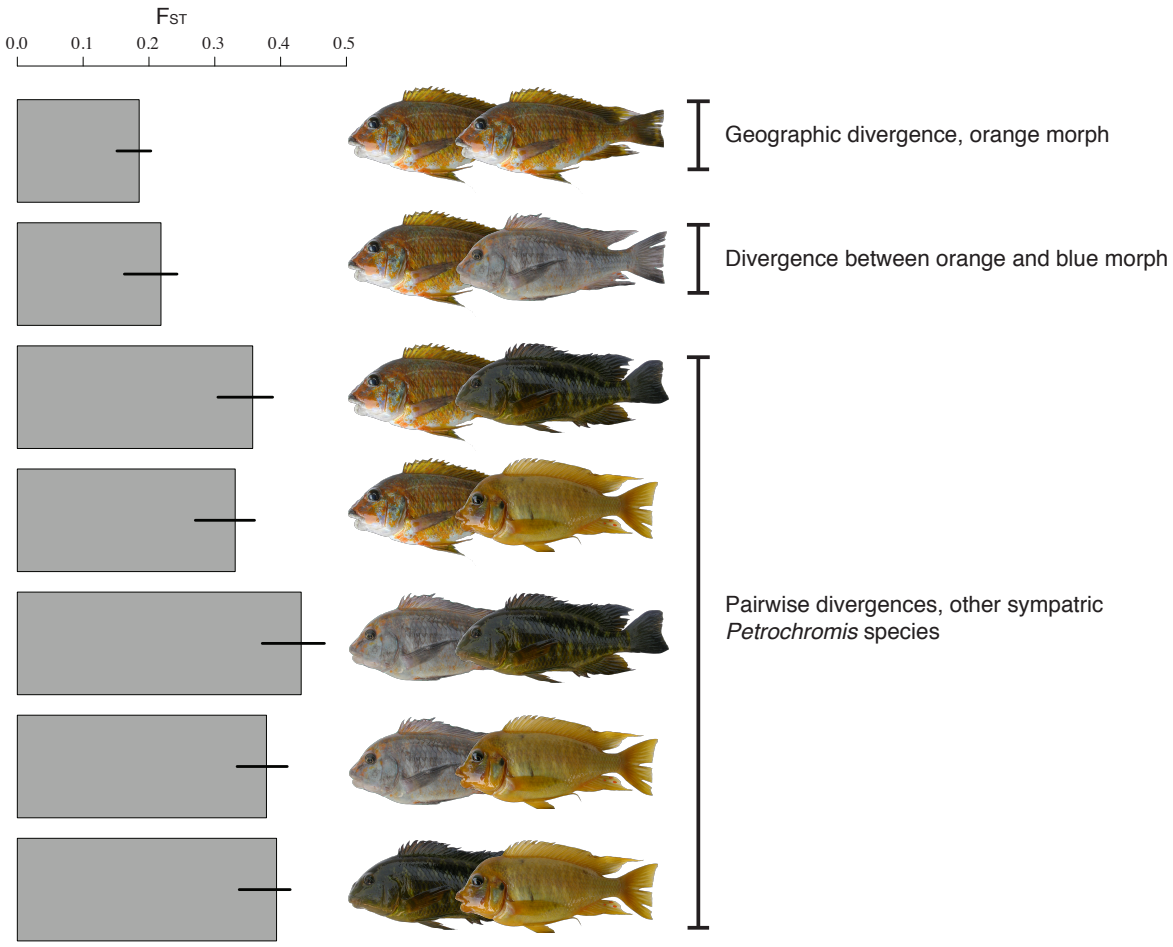


Figure 2.3. Nuclear divergence between the orange and blue *Petrochromis* morphs is of equivalent magnitude to geographic divergence within the orange morph, and substantially smaller than divergence between other sympatric *Petrochromis* morphs. Pairwise F_{ST} values for the northernmost orange morph population with the southernmost (top), the value for pooled orange morph populations vs. blue morph (second from top) and all other pairwise combinations of the orange and blue morphs with *P.* sp. "moshi" and *P.* sp. "green". Black error bars represent minimum and maximum F_{ST} values from jackknifing over loci.

Discussion

We found substantial genetic evidence for very recent divergence between two sympatric cichlid color morphs in Lake Tanganyika. In STRUCTURE analyses, all individuals identified in the field as the blue morph are unambiguously assigned to a genetic group separate from sympatric orange morph individuals. This result implies that mechanisms inhibiting extensive hybridization between the orange and blue *Petrochromis* morphs have evolved, allowing their differentiation to be maintained in sympatry. Pairwise F_{ST} values between the blue morph group and the orange morph group are equivalent in magnitude to those between orange morph populations separated by only 60 km of shoreline, suggesting that the divergence between these sympatric morphs occurred recently. Other comparable pairwise divergences for closely related sympatric *Petrochromis* species are substantially higher than that between the orange and blue morphs (Figure 2.3). This genetic signature of recent speciation has not been previously described in Lake Tanganyikan cichlids, and is unexpected given that these fish lack the hallmark sexual dimorphism and striking color polymorphisms present in rapidly and recently diversifying haplochromine cichlid lineages (Seehausen et al. 1999b).

Recent divergence and evidence for the evolution of barriers to gene flow

Previous phylogenetic work has shown that reciprocal monophyly of Tanganyikan cichlid species at mtDNA loci is effectively universal (Sturmbauer et al. 2003), in strong contrast to extensive sharing of mitochondrial haplotypes among cichlid species in Lakes Malawi and Victoria (e.g. Meyer et al. 1990; Nagl et al. 1998; Won et al. 2006). Unlike the Tanganyikan norm, we find extensive mitochondrial haplotype sharing between the orange and blue

Petrochromis morphs of this study (Figure 2.1b). This lies in distinct contrast to comparisons of the other sympatric *Petrochromis* species found in the Kigoma region, where sequence divergence between species is 2-4% and all sequenced individuals are reciprocally monophyletic at the species level (Figure 2.1a and Appendix 2, Figure S1).

Despite their extensive sharing of mtDNA haplotypes, we find strong genotypic support for the sympatric orange morph and blue morph groups being separately breeding populations. Bayesian analysis of microsatellite variation assigns all fish identified in the field as the blue morph to a single genetic group, with assignments greater than 95% in all cases (Figure 2.2). These fish were collected at the same field sites and in the same nets as orange morph individuals collected from both "north" and "mid" spatial genetic groups. Therefore, regardless of their geographic origin, blue morph individuals are more closely related to each other at nuclear loci than they are to orange morph individuals. This genetic pattern provides strong evidence that barriers to gene flow between the morphs have evolved, allowing them to remain distinct genetic groups in sympatry.

The existence of extensive sharing of mitochondrial haplotypes could be the result of either a) introgression associated with recent and/or ongoing hybridization, b) recent divergence and incomplete lineage sorting at mitochondrial loci, or a combination of these processes. We examine each of these scenarios below.

Hybridization could produce sharing of mitochondrial haplotypes between orange and blue morph *Petrochromis*. However, assignment probabilities of greater than 88% for the 20 of the 23 fish assigned to the blue morph genetic group suggest that ongoing hybridization is not extensive. The few fish that have admixed genetic backgrounds based on STRUCTURE

assignment probabilities could be backcrossed hybrid individuals, but this pattern could also result from sharing of ancestral polymorphism. If hybridization were extensive, many more individuals with admixed genetic backgrounds would be expected**.**

Because extensive ongoing hybridization would not maintain the clear nuclear differentiation that we observe, we interpret the shared mitochondrial haplotypes between orange and blue morphs as most likely the result of incomplete lineage sorting due to the recency of the divergence event creating these morphs. Our data do not exclude the possibility of low levels of ongoing hybridization, but they support a substantial degree of reproductive isolation between the morphs, despite their sympatric distribution. This evidence for reproductive isolation in sympatry suggests that we can consider these morphs to be incipient species.

Speciation scenarios

Given that these *Petrochromis* morphs are currently in sympatry and that their divergence is recent, it is possible that they diverged in sympatry. However, the current sympatric distribution of these fishes does not necessarily imply sympatric speciation, as allo- or parapatric divergence, followed by a return to sympatry, are also possible. *Petrochromis* species have extremely fine-scale spatial genetic structuring, and extensive sandy habitats appear to inhibit gene flow between populations (Wagner and McCune 2009). Additionally, many of the recognized color variants in this and other closely related genera are geographically separated within Lake Tanganyika, implying that geographic divergence in color is common in these fishes (Kohda et al. 1996; Baric et al. 2003). Large-scale geographic surveys of the distribution of these species

would help to assess the possibility of secondary sympatry in the Kigoma region, as their distribution outside of the 60 km of coastline surveyed in this study is unknown.

That said, if sympatry between orange and blue morphs is due to secondary contact, either geographic divergence must have been on small geographic scales, or the morphs underwent a period of hybridization upon secondary contact that decreased their differentiation at nuclear loci. Even short geographic distances separating populations of the orange morph results in divergences of equivalent magnitude to that observed between orange and blue morphs. The scale of geographic divergence in the orange morph does not appear to be unusual for the genus, as sympatric *P.* sp. "moshi" exhibits geographic divergence in the same region that is equivalent in magnitude to that of the orange *Petrochromis* morph (Wagner and McCune 2009). Regardless of the geography of speciation in this scenario, the current sympatric distribution of these *Petrochromis* morphs implies that barriers to gene flow have evolved, and that these barriers act to inhibit extensive gene flow between them. Evidence from studies of hybrid viability suggests that pre-mating isolating barriers are far more important to speciation in other cichlid species than are post-mating barriers (Stelkens et al. 2010). Numerous studies of Lake Malawi and Lake Victoria cichlids support a role for male color in assortative mating (e.g. Allender et al. 2003; Knight and Turner 2004; Maan et al. 2004; Pauers et al. 2004). Three studies using the Tanganyikan tropheine genus *Tropheus* also provide some evidence for assortative mating among color morphs based on coloration (Salzburger et al. 2006; Egger et al. 2008; Egger et al. 2010), although *Tropheus* color morphs are not found in natural sympatry. Given the common role for coloration in assortative mating in cichlids, it is reasonable to suggest

that coloration may play an important role in mate choice in the *Petrochromis* morphs studied here.

A further suggestion that color may be important in assortative mating in this system is that the color differences between these morphs apparently arise late in ontogeny. We collected no juvenile individuals with blue coloration, yet many juveniles not identified as the blue morph in the field are members of the blue morph genetic group (Figure 2.2). This suggests that juveniles are not distinct in color. Late ontogeny color shifts have also been described in Lake Victoria cichlids (Maan et al. 2006) and in neotropical Midas cichlids; in the latter case color morphs show evidence for assortative mating (Barlow 1986; Barlow et al. 1990; Elmer et al. 2009). A correlation between the timing of sexual maturity and the time at which color differences arise in ontogeny suggests a role for color in sexual selection.

Sympatric divergence of the orange and blue *Petrochromis* morphs is also a possibility, given their relative magnitudes of genetic divergence and their current sympatric distribution. Both theoretical and empirical work shows that interactions between ecology and sexual selection may have important impacts on the mechanisms whereby new species arise (Ritchie 2007; Maan and Seehausen 2011), and these linkages may be particularly important in sympatric speciation. Recent work on Lake Victoria cichlids provides support for a "sensory drive" mechanism for speciation, where the light environment, which changes as a function of water depth and clarity, provides a basis for a) natural selection for environment-specific visual acuity and b) sexual selection for male sexual signals, that results in speciation along gradients of intermediate slope (Kawata et al. 2007; Seehausen et al. 2008; Maan and Seehausen 2010). This is an intriguing model to consider in the context of cichlid speciation in other environments. However, Lake

Tanganyika's littoral light environment differs substantially from that of Lake Victoria, and it remains unknown whether cichlid speciation via similar mechanisms could occur in this setting. Ecological factors other than depth could also influence the maintenance of diversity and/or speciation in this group. Environmental heterogeneity contributes to the maintenance of color polymorphism in many animal systems (e.g. Rosenblum 2006; Gray et al. 2008), and if this heterogeneity is stable over time, it can lead to niche partitioning (Endler and Thery 1996; Leal and Fleishman 2002) and in some circumstances could promote speciation (Chunco et al. 2007; Gray and McKinnon 2007). The rocky environments that constitute *Petrochromis* habitat are complex, and different microhabitats might offer consistent differences in light environment that could facilitate natural and sexual selection via sensory drive in a manner analogous to depth gradients. Alternatively, assortative mating could result by habitat sorting alone if color was associated with microhabitat preference, and it could thereby influence the maintenance and/or the origins of diversity. Whether ecological differences arise in the process of speciation or after speciation, and how ecology and sexual signals interact during and after speciation, are compelling and open questions in cichlid diversification and in adaptive radiation in general.

In conclusion, genetic evidence suggests that speciation in the *Petrochromis* morphs studied here is more recent than in any previously studied cichlid species pair in Lake Tanganyika. Although *Petrochromis* and other tropheine cichlids lack the hallmark sexual dichromatism and color polymorphisms present in rapidly diversifying haplochromine cichlid lineages, it is perhaps not a coincidence that we find recent speciation in the Lake Tanganyikan group that is sister to the haplochromines. Comparative studies are key to understanding the origins and maintenance of diversity in ecologically complex and hyper-diverse systems. Understanding the ecological

context for speciation in this Tanganyikan system will provide context for studies of cichlid speciation on similarly recent timescales in the younger lakes that have produced such spectacular examples of vertebrate diversity.

Acknowledgements

This chapter was written in collaboration with Amy R. McCune and Irby Lovette. Many thanks to Pete McIntyre and Ellinor Michel for introducing CEW to this system. Thanks to Rick Harrison, Martine Maan, Dan Rabosky, Matt Carling, and the Lovette lab group for comments on this manuscript, to Laura Stenzler, Chris Makarewich and Amanda Talaba for technical consultation during lab work, to Aurélie Coulon for analytical advice and to George Kazumbe for essential help with fish collection. The Nyanza Project, TAFIRI, the Cornell University Museum of Vertebrates and the Fuller Evolutionary Biology Program provided logistical support. Funding was provided by SSB, ASIH, ACA, Sigma Xi, Cornell EEB and Cornell Mellon grants to CEW, and NSF DDEP grant INT-0724247 to CEW and ARM. In addition, NSF grants ATM-0223920 and DBI-0353765 (Nyanza Project) provided logistical support in the field. Part of this work was carried out using the resources of the CBSU at Cornell University, which is partially funded by Microsoft.

REFERENCES

Allender, C. J., O. Seehausen, M. E. Knight, G. F. Turner, and N. Maclean. 2003. Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. Proc. Natl. Acad. Sci. USA 100:14074-14079.

Baric, S., W. Salzburger, and C. Sturmbauer. 2003. Phylogeography and evolution of the Tanganyikan cichlid genus Tropheus based upon mitochondrial DNA sequences. J. Mol. Evol. 56:54-68.

Barlow, G. W. 1986. Mate choice in the monogamous and polychromatic Midas cichlid, Cichlasoma citrinellum. J. Fish Biol. 29:123-133.

Barlow, G. W., R. C. Francis, and J. V. Baumgartner. 1990. Do the colors of parents, companions and self influence assortative mating in the polychromatic Midas cichlid? Anim. Behav. 40:713-722.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 1996-2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Boulenger, G. A. 1898. Report on the fishes recently obtained by Mr. J. E. S. Moore in Lake Tanganyika. . Pp. 494-497. Proceedings of the General Meetings for Scientific Business of the Zoological Society of London 1898, London.

Brichard, P. 1989. Pierre Brichard's book of cichlids and all the other fish of Lake Tanganyika. T.F.H. Publications, New Jersey.

Chunco, A. J., J. S. McKinnon, and M. R. Servedio. 2007. Microhabitat variation and sexual selection can maintain male color polymorphisms. Evolution 61:2504-2515.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657-1659.

Cohen, A. S., M. Soreghan, and C. Scholz. 1993. Estimating the age of ancient lake basins: An example from Lake Tanganyika. Geology 21:511-514.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.

Day, J. J., J. A. Cotton, and T. G. Barraclough. 2008. Tempo and mode of diversification of Lake Tanganyika cichlid fishes. PLoS ONE 3:e1730.

Dijkstra, P. D., S. van Dijk, T. G. G. Groothuis, M. E. R. Pierotti, and O. Seehausen. 2009. Behavioral dominance between female color morphs of a Lake Victoria cichlid fish. Behav. Ecol. 20:593-600.

Egger, B., S. Koblmuller, C. Sturmbauer, and K. M. Sefc. 2007. Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. BMC Evol. Biol. 7:137.

Egger, B., K. Mattersdorfer, and K. M. Sefc. 2010. Variable discrimination and asymmetric preferences in laboratory tests of reproductive isolation between cichlid colour morphs. J. Evol. Biol. 23:433-439.

Egger, B., B. Obermuller, E. Eigner, C. Sturmbauer, and K. M. Sefc. 2008. Assortative mating preferences between colour morphs of the endemic Lake Tanganyika cichlid genus Tropheus. Hydrobiologia 615:37-48.

Elmer, K. R., T. K. Lehtonen, and A. Meyer. 2009. Color assortative mating contributes to sympatric divergence of neotropical cichlid fish. Evolution 63:2750-2757.

Endler, J. A., and M. Thery. 1996. Interacting effects of lek placement, display behavior, ambient light, and color patterns in three neotropical forest-dwelling birds. Am. Nat. 148:421-452.

Genner, M. J., O. Seehausen, D. F. R. Cleary, M. E. Knight, E. Michel, and G. F. Turner. 2004. How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? J. Biogeogr. 31:93-102.

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

Gray, S. M., and J. S. McKinnon. 2007. Linking color polymorphism maintenance and speciation. Trends Ecol. Evol. 22:71-79.

Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310-2314.

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

Joyce, D. A., D. H. Lunt, M. J. Genner, G. F. Turner, R. Bills, and O. Seehausen. 2010. Repeated colonization and hybridization characterize the Lake Malawi cichlid fish species flock. Curr. Biol. in press.

Kawata, M., A. Shoji, S. Kawamura, and O. Seehausen. 2007. A genetically explicit model of speciation by sensory drive within a continuous population in aquatic environments. BMC Evol. Biol. 7:99.

Knight, M. E., and G. F. Turner. 2004. Laboratory mating trials indicate incipient speciation by sexual selection among populations of the cichlid fish Pseudotropheus zebra from Lake Malawi. Proc. R. Soc. B 271:675-680.

Koblmuller, S., B. Egger, C. Sturmbauer, and K. M. Sefc. 2010. Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. Mol. Phylogen. Evol. 55:318-334.

Kocher, T. D. 2004. Adaptive evolution and explosive speciation: the cichlid fish model. Nature Reviews Genetics 5:288-298.

Kohda, M., Y. Yanagisawa, T. Sato, K. Nakaya, Y. Niimura, K. Matsumoto, and H. Ochi. 1996. Geographical colour variation in cichlid fishes at the southern end of Lake Tanganyika. Environ. Biol. Fishes 45:237-248.

Leal, M., and L. J. Fleishman. 2002. Evidence for habitat partitioning based on adaptation to environmental light in a pair of sympatric lizard species. Proc. R. Soc. B 269:351-359.

Maan, M. E., M. P. Haesler, O. Seehausen, and J. J. M. Van Alphen. 2006. Heritability and heterochrony of polychromatism in a Lake Victoria cichlid fish: Stepping stones for speciation? J. Exp. Zool. B Mol. Dev. Evol. 306B:168-176.

Maan, M. E., and O. Seehausen. 2010. Mechanisms of species divergence through visual adaptation and sexual selection: Perspectives from a cichlid model system. Current Zoology 56:285-299.

Maan, M. E., and O. Seehausen. 2011. Ecology, sexual selection and speciation. Ecol. Lett.

Maan, M. E., O. Seehausen, L. Soderberg, L. Johnson, E. A. P. Ripmeester, H. D. J. Mrosso, M. I. Taylor, T. J. M. van Dooren, and J. J. M. van Alphen. 2004. Intraspecific sexual selection on a speciation trait, male coloration, in the Lake Victoria cichlid Pundamilia nyererei. Proc. R. Soc. B 271:2445-2452.

McCune, A. R. 1997. How fast is speciation: molecular, geological and phylogenetic evidence from adaptive radiations of fishes. Pp. 585-610 *in* T. Givnish, and K. Sytsma, eds. Molecular Evolution and Adaptive Radiation. Cambridge University Press.

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

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

Pauers, M. J., J. S. McKinnon, and T. J. Ehlinger. 2004. Directional sexual selection on chroma and within-pattern colour contrast in Labeotropheus fuelleborni. Proc. R. Soc. B 271:S444-S447.

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

Pritchard, J. K., and W. Wen. 2003. Documentation for STRUCTURE software: Version 2. . Available from http://pritch.bsd.uchicago.edu.

Ritchie, M. G. 2007. Sexual selection and speciation. Annu. Rev. Ecol., Evol. Syst. 38:79-102.

Roberts, R. B., J. R. Ser, and T. D. Kocher. 2009. Sexual Conflict Resolved by Invasion of a Novel Sex Determiner in Lake Malawi Cichlid Fishes. Science 326:998-1001.

Ronquist, F., and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

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

Rosenblum, E. B. 2006. Convergent evolution and divergent selection: Lizards at the White Sands ecotone. Am. Nat. 167:1-15.

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

Salzburger, W., A. Meyer, S. Baric, E. Verheyen, and C. Sturmbauer. 2002. Phylogeny of the Lake Tanganyika cichlid species flock and its relationship to the central and east African Haplochromine cichlid fish faunas. Syst. Biol. 51:113-135.

Salzburger, W., H. Niederstatter, A. Brandstatter, B. Berger, W. Parson, J. Snoeks, and C. Sturmbauer. 2006. Colour-assortative mating among populations of Tropheus moorii, a cichlid fish from Lake Tanganyika, East Africa. Proc. R. Soc. B 273:257-266.

Seehausen, O., E. Koetsier, M. V. Schneider, L. J. Chapman, C. A. Chapman, M. E. Knight, G. F. Turner, J. J. M. van Alphen, and R. Bills. 2003. Nuclear markers reveal unexpected genetic variation and a Congolese-Nilotic origin of the Lake Victoria cichlid species flock. Proc. R. Soc. B 270:129-137.

Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V. Schneider, M. E. Maan, H. Tachida, H. Imai, and N. Okada. 2008. Speciation through sensory drive in cichlid fish. Nature 455:620-U623.

Seehausen, O., J. J. M. Van Alpen, and F. Witte. 1999a. Can ancient colour polymorphisms explain why some cichlid lineages speciate rapidly under disruptive sexual selection? Belg. J. Zool. 129:43-60.

Seehausen, O., J. J. M. van Alphen, and R. Lande. 1999b. Color polymorphism and sex ratio distortion in a cichlid fish as an incipient stage in sympatric speciation by sexual selection. Ecol. Lett. 2:367-378.

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

Seutin, G., and B. N. White. 1991. Preservation of avian blood and tissue samples for DNA analysis. Can. J. Zool. 69:82-90.

Stelkens, R. B., K. A. Young, and O. Seehausen. 2010. The accumulation of reproductive incompatibilities in African cichlid fish. Evolution 64:617-632.

Sturmbauer, C., U. Hainz, S. Baric, E. Verheyen, and W. Salzburger. 2003. Evolution of the tribe Tropheini from Lake Tanganyika: synchronized explosive speciation producing multiple evolutionary parallelism. Hydrobiologia 500:51-64.

Wagner, C. E., and A. R. McCune. 2009. Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. Evolution 63:1312-1326.

Wagner, C. E., P. B. McIntyre, K. S. Buels, D. M. Gilbert, and E. Michel. 2009. Diet predicts intestine length in Lake Tanganyika's cichlid fishes. Funct. Ecol. 23:1122-1131.

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

Won, Y. J., Y. Wang, A. Sivasundar, J. Raincrow, and J. Hey. 2006. Nuclear gene variation and molecular dating of the cichlid species flock of Lake Malawi. Mol. Biol. Evol. 23:828-837.

CHAPTER 3

ECOLOGICAL OPPORTUNITY AND SEXUAL SELECTION TOGETHER PREDICT ADAPTIVE RADIATION

Summary

Adaptive radiations are a major feature in the evolution of biodiversity, but explaining why some clades undergo adaptive radiations, diversifying extensively into many and varied species, while others do not, remains a fundamental challenge. Evolutionary diversification is thought to be influenced by both intrinsic lineage-specific traits and extrinsic environmental factors (Simpson 1953; Ricklefs 2007; Barraclough 2010), yet few studies have successfully untangled their interactions. We take advantage of a broad-scale "natural experiment" in adaptive radiation involving cichlid fishes in African lakes. The spectacular cichlid radiations of the African great lakes are well-known, but cichlids have also radiated within more than thirty other African lakes. Furthermore, more than 70 cichlid lineages have colonized lakes but not diversified. These replicated events allow us to examine the factors that determine when adaptive radiation happens and when it does not. We compiled data on cichlid colonization and diversification in 46 African lakes, along with environmental variables and information about the traits of colonizing cichlid lineages. We find that lineage-specific traits related to sexual selection and environmental factors related to ecological opportunity both strongly influence whether cichlids radiate. Among extrinsic factors, lake depth, net solar radiation, and time for diversification are positively associated with diversification. Negative associations between diversification and lake surface

area indicate that cichlid speciation is not limited by area, as it is for terrestrial taxa (Kisel and Barraclough 2010). Among intrinsic traits, sexual dichromatism, a surrogate for the intensity of sexual selection, is consistently positively associated with diversification. Our results indicate that, for cichlids, it is the coincidence between environmental opportunity and sexual selection that best predicts whether adaptive radiation will occur.

Adaptive radiations are iconic systems for the study of evolutionary processes. The rapidity of speciation and the wealth of ecological diversity, particularly within geologically young adaptive radiations, have greatly advanced our understanding of the process of biological diversification (Schluter 2000; Losos 2010). However, for many examples of adaptive radiation, there are closely related lineages that have not diversified. In other cases, some habitats or geographical regions are a dramatic theatre for diversification while others remain depauperate. What factors determine whether a lineage diversifies upon entry into a habitat? Why is it that some lineages diversify dramatically, whereas closely related lineages in the same habitat do not?

On one hand, we might view adaptive radiation as a consequence of ecological opportunity (Simpson 1953; Schluter 2000). Extrinsic factors related to ecological opportunity that have been linked to adaptive radiation include a paucity of competing lineages (Simpson 1953; Losos 2010), predation regime (Vamosi 2003), biotic insularity (Rosindell and Phillimore 2011), habitat complexity (Price et al. 2011), and habitat area (Kisel and Barraclough 2010). Latitude (Mittelbach et al. 2007) and energy, measured as solar radiation or primary productivity (Evans et al. 2005), have been linked to variation in broad-scale patterns of diversity (e.g. the latitudinal

diversity gradient), but these factors have not been previously investigated in the context of geographically-circumscribed adaptive radiations. On the other hand, differences in diversification may result from differences in lineage-specific traits. Traits hypothesized to be linked to speciation rates include factors such as intensity of sexual selection (Kraaijeveld et al. 2010), ecological specialization (Farrell 1998), ecological versatility (Liem 1973) and spatial vagility (Kisel and Barraclough 2010). Although this approach has successfully linked species traits to diversification rates in a number of taxa, the overall proportion of variation explained is generally low (Ricklefs 2007). A major conceptual challenge to future work on diversification and its causes lies in identifying the relative roles of intrinsic and extrinsic factors, and how the importance of these factors and their interactions varies among taxa and environments.

Both environmental factors and lineage-specific traits influence the evolution and distribution of biodiversity at broad macroecological scales (Ricklefs 2007; Barraclough 2010), but rarely, if ever, have the influence of multiple extrinsic and intrinsic factors, and their interactions, been considered simultaneously in the study of geographically-circumscribed adaptive radiation. Since the discovery of the species-rich African lake cichlid faunas, hypotheses invoking intrinsic traits (Liem 1973; Seehausen and van Alphen 1999), environmental factors (Fryer 1959; Sturmbauer et al. 2001), and their interactions (Fryer 1959; Seehausen 2007) as explanations for the spectacular diversity of these fishes have proliferated, yet these hypotheses remained untested at macroevolutionary scales in an explicit phylogenetic context. Furthermore, while most efforts have focused on the three major radiations in Lakes Victoria, Tanganyika, and Malawi, cichlids have independently diversified within African lakes on more than 30 occasions. Even more important for testing the factors that drive diversification is that on more than 70 occasions

cichlid lineages have entered lakes but have *not* diversified. These replicated cases of intralacustrine diversification, and failure to diversify, provide a powerful opportunity to test which factors predict whether a cichlid population will diversify or not. At one extreme, diversification could be entirely environmentally determined; in this case, lake environment alone would predict the occurrence of speciation. At the other extreme, lake environment could be unimportant compared to the influence of colonizing species' traits; traits alone could predict the occurrence of speciation. Between these extremes, diversification could depend on both environmental and lineage-specific factors and their interactions. Finally, the occurrence of intralacustrine diversification could be entirely stochastic, and in this case neither environmental variables nor intrinsic traits would predict diversification.

We used nuclear and mitochondrial sequence data to build a multi-gene molecular phylogeny for African cichlids (see Methods, Appendix 3 for details), and phylogenetically placed all lacustrine occurrences of African cichlids (Figure 3.1, also see Methods, Appendix 3) on this tree. We then collated information on lake characteristics for 46 lakes harboring cichlid lineages across the African continent (Figure 3.1; Appendix 3**)**. Environmental factors included lake depth, net solar radiation (the difference between the influx of solar energy and that reflected back into the atmosphere at a given geographic location), latitude, elevation, and the presence or absence of large predatory fish. We also calculated "time for diversification" for each colonization, as either the time since last evidence for lake desiccation, or the stem age of the lineage as calculated from calibrated molecular phylogenies (see Methods, Appendix 3). We collected data on the intrinsic traits of cichlid lineages, including the presence of a polygamous mating system, the presence of mouthbrooding as a parental care strategy, the presence of generalized egg

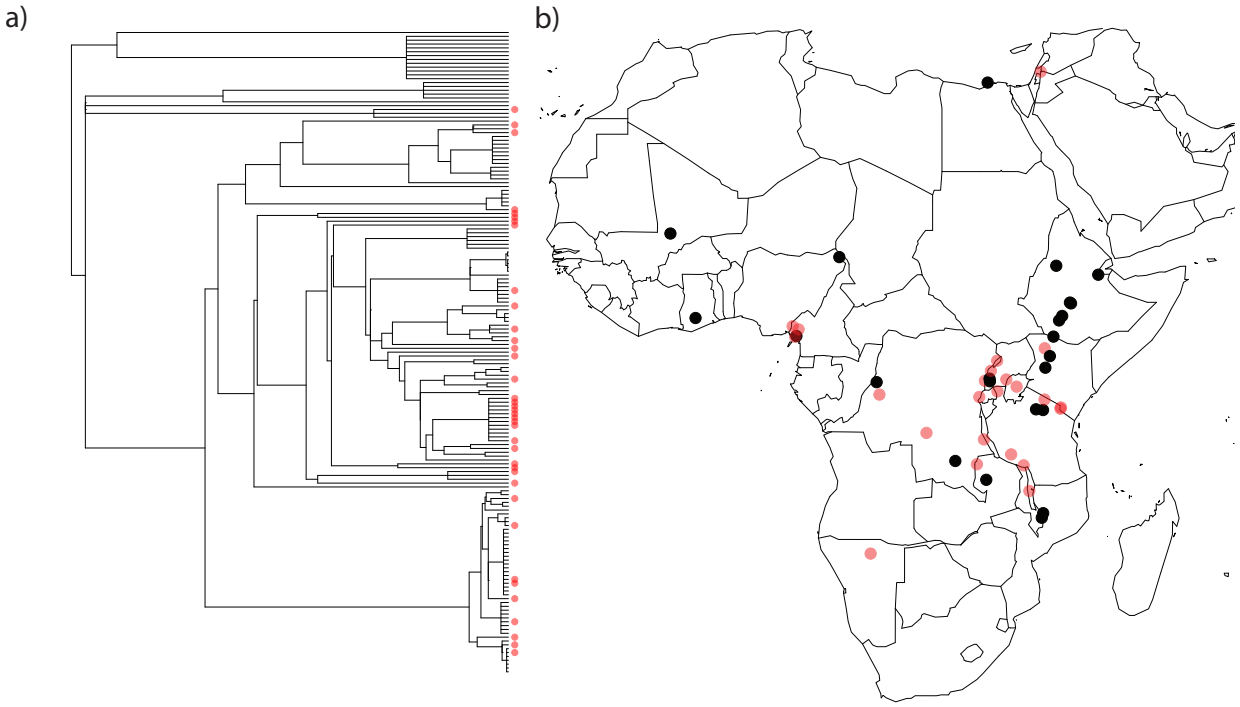


Figure 3.1. Cichlid diversification is phylogenetically and geographically widespread. a) The distribution of intralacustrine diversification across the African cichlid phylogeny. Each tip represents one lineage in a lake; red dots indicate diversifying lineages (lineages with at least one intralacustrine speciation event). b) The geographic distribution of cichlid diversification in lakes across Africa. Each dot represents a cichlid lineage present within a lake. Black dots indicate no endemic diversification; red dots indicate at least one intralacustrine speciation event.

dummies and/or the morphologically derived-type "haplochromine" egg dummies (Greenwood 1979) on the anal fin of the males (used in fertilization of eggs in the mouth of the female), and the presence of strong sexual dichromatism. Each of these traits has been hypothesized to be linked to sexual selection and/or intensity of sexual selection in cichlids, which case studies show is important to cichlid speciation (Seehausen et al. 2008). We used model averaging approaches and phylogenetic logistic regression to test for associations between these predictor variables and cichlid "diversification state" – that is, whether a lineage has diversified upon entering a lake or has failed to do so (see Methods). This analytical approach identifies factors

associated with diversification, and these factors may be either causally related to diversification, or their association could arise through correlation with other unmeasured causal factors.

We find that environmental factors linked to ecological opportunity and lineage-specific traits related to the strength of sexual selection significantly predict cichlid diversification in African lakes. In initial single-predictor variable analyses, the strongest environmental predictors of diversification were lake depth and time for diversification (Appendix 3, Supplementary Table 6). Among intrinsic traits, sexual dichromatism was the strongest predictor of diversification.

We examined the combined influence of predictor variables on diversification state in multiple regression models using AICc-based model averaging (Burnham and Anderson 2002) followed by phylogenetic multiple logistic regression of a reduced predictor variable set (see Methods). In cases where predictor variables were highly correlated (e.g. lake depth and age, net solar radiation and latitude), we either eliminated one of the two correlated predictors or used various strategies to parse out the effects of the two variables (see Methods, Appendix 3 for details). The best-supported predictor variables in our multiple regression models include both environmental variables and lineage-specific traits (Table 3.1). We consider variables with relative importance (RI) scores greater than 0.7 that are also significant in our phylogenetically controlled models as well-supported predictors of diversification. Well-supported predictor variables included lake depth, net solar radiation and sexual dichromatism, all of which were positively related to diversification, and lake surface area, which had a negative relationship with diversification (for all RI > 0.90). These variables remained strong predictors when we removed Lake Tanganyika, an outlier in both age and depth, from the analyses (see Appendix 3). In analyses excluding Lake Tanganyika, the same three environmental variables were the strongest extrinsic factors

associated with diversification (RI > 0.95), and a negative effect of elevation also emerged as a strong predictor (RI = 0.88). Sexual dichromatism remained the strongest intrinsic trait predictor of diversification. Although there were associations between haplochromine-type egg dummies and diversification in model-averaged results, this result was not significant when we accounted for phylogeny (Table 3.1).

Table 1. Multiple logistic regression models reveal that environmental factors and lineage-specific traits together best explain cichlid diversification in African lakes. Among environmental variables, there are positive associations between cichlid diversification and lake depth, environmental energy, and lake surface area is a negative predictor of diversification. Among lineage-specific traits, the presence of sexual dichromatism is a significant predictor of diversification. These results are consistent when we exclude lineages present in Lake Tanganyika, an outlier in terms of depth, from the analysis.

	Full Dataset							Excluding Lake Tanganyika						
	nonphylogenetic			phylogenetic				nonphylogenetic			phylogenetic			
Predictor	Relative-importance value	Estimate	± SE	Estimate	±SE	Wald Z	p-value	Relative-importance value	Estimate	± SE	Estimate	±SE	Wald Z	p-value
Lake Surface Area	**0.989**	-0.458	0.156	-0.421	0.128	**-3.283**	**0.001**	**0.984**	-0.471	0.167	-0.396	0.135	**-2.941**	**0.002**
Lake Depth	0.262	0.044	0.075					0.461	0.185	0.126				
Energy	**0.999**	0.117	0.036	0.094	0.029	**3.239**	**0.001**	**0.999**	0.131	0.040	0.108	0.032	**3.379**	**0.000**
Residual Latitude	0.473	-0.072	0.049					0.554	-0.102	0.063				
Elevation	**0.877**	-0.003	0.001	-0.002	0.001	**-1.966**	**0.025**	**0.952**	-0.004	0.002	-0.004	0.001	**-3.043**	**0.001**
Predators	0.307	0.230	0.269					0.301	0.227	0.297				
Polygamous Mating System	0.249	-0.050	0.276					0.320	0.597	0.702				
Egg dummies	0.468	-0.710	0.521					0.402	-0.656	0.578				
Haplo egg dummies	**0.901**	1.997	0.817	1.114	0.761	1.463	0.072	0.945	2.569	1.026	1.945	0.797	**2.440**	**0.007**
Sexual Dichromatism	**0.893**	1.869	0.770	1.484	0.746	**1.989**	**0.023**	0.895	2.305	0.946	1.882	0.883	**2.130**	**0.017**
Time	**0.988**	0.660	0.177	0.605	0.147	**4.128**	**0.000**	0.580	0.312	0.177				

Both single-predictor and multiple regression models reveal a strong association between lake depth and cichlid diversification. This may be in part due to increased environmental stability through time afforded by increased lake depth, and indicate that lineages are more likely to diversify when given more time to do so. In addition, increased depth also increases habitat availability and dimensionality for fish, and this could positively influence diversification. Depth

partitioning of resources and reproduction appears to be important in many cases of intralacustrine speciation in fishes, and spawning depth is often a major ecological difference among closely related fish species radiations (Vonlanthen et al. 2009; Ingram 2011). Case studies indicate that depth-specific spawning segregation and ecological adaptation can be key factors in cichlid speciation (Seehausen et al. 2008), and our finding here is consistent with these mechanisms being important for predicting macroevolutionary patterns of cichlid diversification.

Net solar radiation emerges as a strong predictor of cichlid diversification in multiple regression models. Links between energy and evolutionary diversification have been frequently hypothesized in the context of latitudinal gradients in species richness (Mittelbach et al. 2007), although only rarely has this relationship been tested in an explicitly evolutionary framework, and it has not been previously investigated in the context of "insular" adaptive radiation. A variety of mechanisms have been proposed to explain relationships between energy and species richness, and those linked to diversification include 1) increased environmental carrying capacity with increased energy availability, hence larger total population size, which may result in increased rates of speciation and/or lower rates of extinction; 2) increased mutation rates and/or shortened generation times in high-energy environments, resulting in increased rates of population differentiation and speciation (Evans et al. 2005). The association between intralacustrine diversification and net solar radiation in cichlids could be linked to diversification via either or both of these mechanisms.

The negative association between diversification and lake surface area in our models shows that in contrast to diversification in terrestrial systems (Kisel and Barraclough 2010), larger areas do not increase the likelihood that colonizing lineages will undergo speciation. Sampling bias could

influence this result: data on species present in very small lakes is rare compared to that for large lakes, and small lakes included in the dataset are frequently those known for their endemic cichlids (e.g. Cameroonian crater lakes; Guineas Sink Hole; see Appendix 3). However, in other systems like *Anolis* lizards on islands, very small areas never house adaptive radiations (Losos and Schluter 2000). Regardless of potential size-related sampling bias, this finding reveals a major contrast between cichlids and terrestrial taxa in that in-situ speciation is not limited by area.

The consistent association between sexual dichromatism and diversification in both our multiple regression models and single-predictor models (Table 3.1, Appendix 3) suggests that the intensity of sexual selection may play a key role in determining whether lineages radiate when given ecological opportunity. Sexual dichromatism is a common proxy for strength of sexual selection in studies of diversification (Kraaijeveld et al. 2010). Divergence among populations in traits under sexual selection, and associated preferences, can readily lead to pre-mating isolation among populations, and thereby facilitate speciation (Lande 1981). Although sexual selection is known to be important in cichlid speciation from case studies, we here show an association between sexual selection and diversification in cichlids at macroevolutionary scales. Examination of the co-occurrence between dichromatism and mating system reveals that sexual dichromatism only evolves in lineages that have polygamous mating systems (Figure 3.3), a pattern predicted if mating system determines opportunity for sexual selection (Trivers 1972). This result indicates that dichromatism is a more direct indicator of the actual strength of sexual selection than is mating system, a pattern that has been shown in birds and other taxa (Kraaijeveld et al. 2010) but had never been tested in cichlids or any other adaptive radiation.

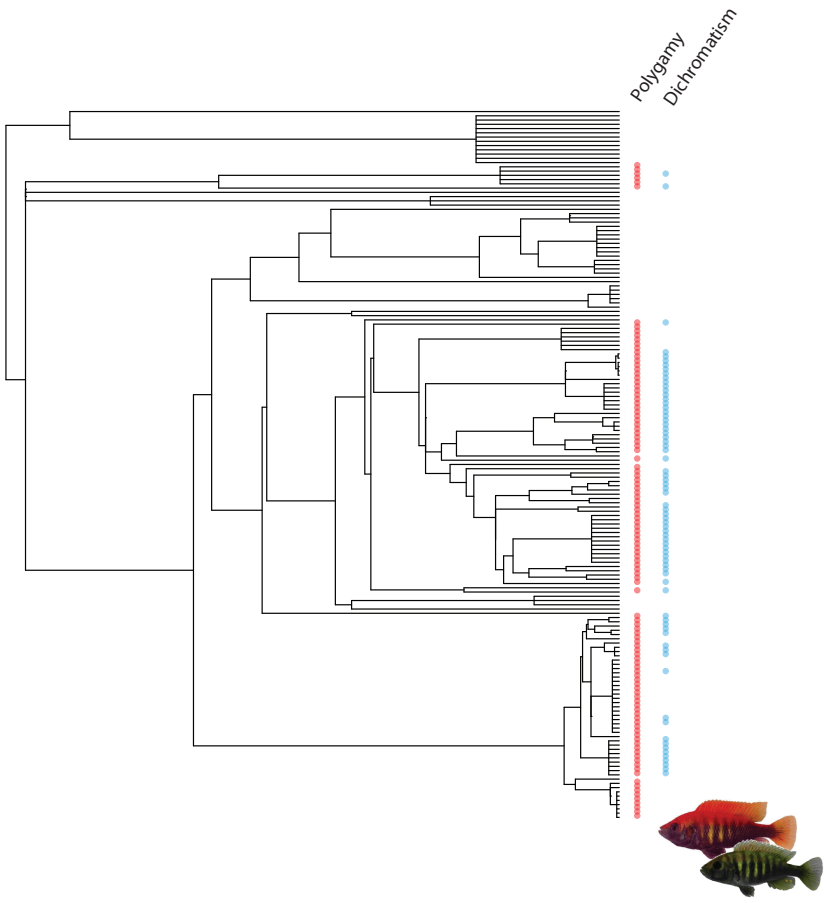


Figure 3.3. All lineages with sexual dichromatism have polygamous mating systems, implying that the evolution of sexual dichromatism only occurs in lineages with polygamous mating systems, and that polygamy is a prerequisite of strong sexual selection in cichlids.

Although African cichlid fishes are an iconic example of adaptive radiation, our analysis shows that there is great heterogeneity in adaptive radiation across this clade–many lineages present in lakes do not diversify. Our results makes clear that high diversification propensity is not an intrinsic property of all cichlids, but one that has evolved in some branches of the cichlid tree. We show here, for the first time, that cichlid adaptive radiations are not a simple function of any one predictor variable, but instead are best predicted by variables representing both extrinsic environmental effects and intrinsic, lineage-specific traits. For cichlids, it is the combined effects

of the intensity of sexual selection, and environmental opportunity, in the form of lake age, size and energy availability, that best predicts whether adaptive radiation will occur. More generally, the multicausal propensity for adaptive radiation helps to explain why only some taxa radiate, even in environmental settings – like islands and lakes – that favor high rates of diversification.

Methods

Phylogenetic framework

We compiled sequence data for nine genes and 656 African cichlid species, with the goal of phylogenetically placing all African cichlid lineages present in lakes. For information about genes, and genbank numbers see Appendix 4. The aligned, concatenated dataset included a total of 6947 base pairs.

We used a maximum likelihood approach in RAxML for phylogenetic analyses (Stamatakis 2006) (see supplement for details). To account for phylogenetic uncertainty, we used 100 replicates of RAxML's rapid bootstrap algorithm and estimated branch lengths for each of these bootstrap replicate topologies. To ultrametricize and time-calibrate this set of trees, we used PATHd8 (Britton et al. 2007). We used four geological dates to time-calibrate the trees: two dates associated with the breakup of Gondwana (the African-Madagascar split and the Madagascar-India split), the age of the earliest known fossil *Oreochromis*, and the age of Lake Nabugabo (see supplement for details). We then drew 95% confidence intervals on node ages from the distribution of branching times estimated from this set of calibrated ultrametric trees.

Cichlid radiation data, ecological variables and species traits

We compiled information about presence of cichlid lineages in lakes across Africa, and the endemic diversity of the lineages present in each of these lakes (see Appendix 3, Supplementary Table 3).

We coded each lineage in each lake as either "diversifying" or "nondiversifying". We identified diversifying lineages as those that had undergone at least one intralacustrine speciation event. This included any lineage that had at least one endemic species in a lake co-occurring with its sister taxon (either a widespread species or a lake endemic itself). Single endemic species not co-occurring with a sister taxon were not considered to be diversifying. We additionally conducted analyses using species richness thresholds of 3 and 5 endemic species to code a lineage as "diversifying" (see Appendix 3). These analyses produced similar results to those presented in the main text.

We compiled information about lineage-level character states for traits potentially linked to cichlid diversification. These included the presence of a polygamous mating system, the presence of mouthbrooding, the presence of generalized egg dummies and specialized haplochromine-type egg dummies on the anal fin of male fish, and the presence of strong sexual dichromatism (see Appendix 3, Supplementary Table 3). Very few of these traits are polymorphic within cichlid lineages. These few instances were coded as missing data for our analyses.

We compiled information on physical and environmental variables for all lakes in the dataset. These include surface area, maximum depth, latitude, net solar radiation and elevation (see Appendix 3, Supplementary Table 2). We chose these variables as the major factors correlating

with lake type, habitat availability, and climate that were available for a large number of lakes. As an additional environmental variable, we included the presence of large predatory fish (genera *Lates*, *Hydrocyanus, Hepsetus*) due to their hypothesized influence on cichlid diversification (Worthington and Ricardo 1936; Fryer 1959).

We calculated maximum time for diversification for lineages using either the midpoint of geological age estimates for the lake (either basin age or most recent desiccation age) or the mean stem age of the radiating group estimated from our calibrated molecular phylogenies. We also conducted analyses using only geological lake ages, and these produced very similar results (see Appendix 3).

Logistic regression models

To account for phylogeny in regression models, we first trimmed the best maximum likelihood topology to include a) only lineages that occur in lakes, b) a single taxon for each lake in which cichlids have diversified. For lineages present in multiple lakes, we added a tip to the tree for each instance where the lineage is found in a unique lake, such that each lineage found in multiple lakes is represented as a polytomy with a tip corresponding to each lake where it is present. We set branch lengths on these added tips to have a total length that matched that expected under a pure birth model (see supplement for details). Using this approach, our trimmed and manipulated phylogenies had a branch for each "opportunity" to diversify – each instance a lineage entered a new lake.

We used phylogenetic logistic regression (Ives and Garland 2010) to assess the relationship between single predictor variables and diversification state. We then used multiple logistic

regression models to assess the combined influence of our predictor variables on cichlid diversification state. Before including the predictor variables in multiple regression models, we checked for collinearity between both continuous and binary predictor variables. We calculated Pearson correlation coefficients (r^2) for all pairs of continuous predictor variables. For binary predictor variables, we used the r^2 equivalent (Menard 2000), r^2_L, as an assessment of collinearity (see supplement for details). We removed one variable from each pair of predictor variables with r^2 (or r^2_L) of greater than 0.6 after preliminary models including variables with correlations higher than this caused analytical problems (inflations of standard error in parameter estimation, a diagnostic of collinearity problems in logistic regression (Quinn and Keough 2002)).

Because we discovered a strong correlation between lake depth and time to diversify during collinearity tests ($r^2 = 0.76$), we conducted further tests to elucidate the relative effects of time and depth. We excluded lakes greater than 150 meters in depth, leaving the remaining data subset largely uncorrelated in time and depth ($r^2 = 0.25$). We compared AIC values among models incorporating time, time + depth, and depth as predictors of cichlid diversification.

We examined the combined influence of predictor variables on diversification state in multiple regression models using a two-step approach. First, we used AICc-based model averaging (Burnham and Anderson 2002) to evaluate the parameter estimates and the relative importance of predictor variables in a likelihood-based framework. We calculated model-averaged parameter estimates and standard errors for each predictor variable using relative AICc weights of models in which the variables appeared. We calculated the relative importance (RI) of each predictor variable as the sum of the AICc-weights of the models including each variable. Second, we included predictor variables with RI values above 0.75 in phylogenetic logistic regression (Ives

and Garland 2010) models to attain phylogenetically corrected regression parameter estimates. We used this two-step approach because likelihood-based phylogenetic logistic regression methods are not available.

Acknowledgements

This chapter was written in collaboration with Ole Seehausen and Luke Harmon. We thank A. Ives for analytical advice, P. McIntyre for access to global energy data, and A. McCune, D. Rabosky, R. Harrison, I. Lovette, E. Michel and G. Mittelbach for discussions and comments on the manuscript. This work was funded by the Swiss National Science Foundation and EAWAG.

REFERENCES

Barraclough, T. G. 2010. Evolving entities: towards a unified framework for understanding diversity at the species and higher levels. Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci. 365:1801-1813.

Britton, T., C. L. Anderson, D. Jacquet, S. Lundqvist, and K. Bremer. 2007. Estimating divergence times in large phylogenetic trees. Syst. Biol. 56:741-752.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference. Springer, New York.

Evans, K. L., P. H. Warren, and K. J. Gaston. 2005. Species-energy relationships at the macroecological scale: a review of the mechanisms. Biological Reviews 80:1-25.

Farrell, B. D. 1998. "Inordinate fondness" explained: Why are there so many beetles? Science 281:555-559.

Fryer, G. 1959. Some aspects of evolution in Lake Nyasa. Evolution 13:440-451.

Greenwood, P. H. 1979. Towards a phyletic classification of the 'genus' Haplochromis (Pisces, Cichlidae) and related taxa. Part 1. Bulletin of the British Museum (Natural History) Zoology 35:265-322.

Ingram, T. 2011. Speciation along a depth gradient in a marine adaptive radiation. Proc. R. Soc. B 278:613-618.

Ives, A. R., and T. Garland. 2010. Phylogenetic Logistic Regression for Binary Dependent Variables. Syst. Biol. 59:9-26.

Kisel, Y., and T. G. Barraclough. 2010. Speciation Has a Spatial Scale That Depends on Levels of Gene Flow. American Naturalist 175:316-334.

Kraaijeveld, K., F. J. L. Kraaijeveld-Smit, and M. E. Maan. 2010. Sexual selection and speciation: the comparative evidence revisited. Biol. Rev. Camb. Philos. Soc.

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

Liem, K. F. 1973. Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Systematic Zoology 22:425-441.

Losos, J. B. 2010. Adaptive Radiation, Ecological Opportunity, and Evolutionary Determinism. American Naturalist 175:623-639.

Losos, J. B., and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. Nature 408:847-850.

Menard, S. 2000. Coefficients of determination for multiple logistic regression analysis. American Statistician 54:17-24.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, A. H. Hurlbert, N. Knowlton, H. A. Lessios, C. M. McCain, A. R. McCune, L. A. McDade, M. A. McPeek, T. J. Near, T. D. Price, R. E. Ricklefs, K. Roy, D. F. Sax, D. Schluter, J. M. Sobel, and M. Turelli. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecol. Lett. 10:315-331.

Price, S. A., R. Holzman, T. J. Near, and P. C. Wainwright. 2011. Coral reefs promote the evolution of morphological diversity and ecological novelty in labrid fishes. Ecol. Lett. 14:462-469.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Experimental design and data analysis for biologists:i-xvii, 1-537.

Ricklefs, R. E. 2007. History and diversity: Explorations at the intersection of ecology and evolution. Am. Nat. 170:S56-S70.

Rosindell, J., and A. B. Phillimore. 2011. A unified model of island biogeography sheds light on the zone of radiation. Ecol. Lett.

Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

Seehausen, O. 2007. Evolution and ecological theory - Chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. Heredity 99:361-363.

Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V. Schneider, M. E. Maan, H. Tachida, H. Imai, and N. Okada. 2008. Speciation through sensory drive in cichlid fish. Nature 455:620-U623.

Seehausen, O., and J. M. van Alphen. 1999. Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? Ecology Letters 2:262-271.

Simpson, G. G. 1953. The major features of evolution. Columbia Univ. Biol. Ser. No. 17:i-xx, 1-434.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.

Sturmbauer, C., S. Baric, W. Salzburger, L. Ruber, and E. Verheyen. 2001. Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. Mol. Biol. Evol. 18:144-154.

Trivers, R. L. 1972. Parental investment and sexual selection. Pp. 136-179 *in* B. Campbell, ed. Sexual Selection and the Descent of Man. Aldine, Chicago.

Vamosi, S. M. 2003. The presence of other fish species affects speciation in threespine sticklebacks. Evolutionary Ecology Research 5:717-730.

Vonlanthen, P., D. Roy, A. G. Hudson, C. R. Largiader, D. Bittner, and O. Seehausen. 2009. Divergence along a steep ecological gradient in lake whitefish (Coregonus sp.). J. Evol. Biol. 22:498-514.

Worthington, E. B., and C. K. Ricardo. 1936. The fish of Lake Tanganyika (other than Cichlidae). Proc Zool Soc London 1936:1061-1112.

CHAPTER 4

ECOLOGICAL CORRELATES OF CICHLID SPECIES RICHNESS IN AFRICAN LAKES

Abstract

A positive relationship between species richness and island size is thought to emerge from an equilibrium between rates of immigration and extinction, but the influence of species diversification on the form this relationship is poorly understood. Here, we show that lake area is a strong predictor of species richness in African cichlid fishes, but only above a distinct threshold in lake size (~1,030 km^2). Contrary to other systems where an area threshold has been observed, this threshold cannot be explained by the onset of in situ speciation in lakes above a certain size, as speciation in cichlids is common in small lakes. To investigate other influences on species richness, we explore the relationship between species richness and energy (measured as net solar radiation), lake depth, lake age, and clade age. We show that total species richness per lake is correlated with measures of lake area, lake depth, and energy, suggesting that these environmental variables are predictors of diversity for the cichlid faunas of African lakes. Additionally, we show that the species richness of clades with in situ diversification (within-lake radiations) is limited by environmental variables, suggesting ecological limits on the species richness of cichlid radiations. We conclude that ecological carrying capacities exist which render the total diversity of cichlids predictable within these lakes, but that these diversities are achieved by lineage-specific diversification outcomes, and thereby producing marked differences in the faunal composition of different lakes.

Introduction

Understanding the factors influencing geographical patterns of species richness is a fundamental goal of ecological research. Perhaps the two most general patterns in spatial ecology are the species-area relationship and the equatorial-polar species richness gradient (Gaston 2000). The species-area relationship became one of the pillars of island biogeography theory (MacArthur and Wilson 1963), one of the most influential theoretical theories in ecology. Hundreds, if not thousands, of studies have shown positive relationships between area and species richness, with characteristic slopes (Losos and Ricklefs 2010). In the context of the latitudinal gradient in species diversity, studies of the relationship between species richness and energy availability are numerous (reviewed in Evans et al. 2005), and positive species-energy relationships are nearly as ubiquitous as are positive species-area relationships (Gaston 2000). However, despite compelling calls to unite these theories about how energy and area influence species richness (Wright 1983), few studies have combined species-energy and species-area theories in the examination of species richness patterns (but see Storch et al. 2005; Hurlbert and Jetz 2010), especially in island systems. Furthermore, even though islands are home to many of evolution's most iconic adaptive radiations, the role of size and energy as predictors of in situ speciation, and the impact of in situ speciation on species richness patterns in these systems, have received little attention.

In classic island biogeography theory, MacArthur and Wilson viewed species richness as the equilibrium between immigration and extinction rates, and treated newly evolved species as equivalent to new immigrants to the system. The implicit assumption is that that species derived from these different sources would not influence equilibrium diversity in distinct ways. However, evidence suggests that speciation within island systems results in a steeper rate of increase in

species richness with area than in communities assembled purely by dispersal (Losos and Schluter 2000; Triantis et al. 2008; Rabosky and Glor 2010; Rosindell and Phillimore 2011). Furthermore, regions wherein the majority of species have arisen in situ ("biological provinces") are well known to have steeper species-area relationships than regions in which immigration and emigration play a proportionally larger role in determining species richness (Rosenzweig 1995; Rosenzweig 2001). Several mechanisms have been proposed to explain this increase in the slope of the species-area relationship. These include both nonequilibrium explanations, such as increased diversification rates in larger areas (Losos and Schluter 2000), and explanations invoking equilibrium diversities, including the idea that the speciation-extinction equilibrium scales with area (Rosenzweig 1992; Rosenzweig 1995; Rosenzweig 2001) and that evolutionarily assembled biotas have greater carrying capacities than biotas assembled via dispersal (Rabosky and Glor 2010).

Species-energy relationships are rarely considered in the context of island biogeography, and, conversely, studies of species-area relationships at latitudinal-diversity-gradient scales have been limited (Gaston 2000; but see Mittelbach et al. 2007). Wright (1983) attempted to unite these ideas, suggesting that the product of energy and area (total energy, hereafter E_{tot}), could replace area in MacArthur and Wilson's (1963) island biogeography models, under the logic that it more succinctly, and more generally, approximates the resource availability within a given area[2]. Although Wright's (1983) paper has been cited hundreds of times, most studies citing it examine species-energy relationships while holding area constant. However, because the rate of species

[2] Although note that MacArthur and Wilson believed that extinction rates, not resource availability, were responsible for the species-area pattern

accrual with area might diverge from that with energy, these two factors might interact to determine patterns of species richness (Hurlbert and Jetz 2010), especially in situations when there is large variation in both area and energy within the study region.

Where island communities are not in equilibrium, the timescale over which we consider species-area and species-energy relationships will be important for understanding species richness patterns if island age impacts richness, either through ecological processes (e.g. community assembly through colonization and extinction) or through evolutionary dynamics (e.g. speciation). Time, measured as either island age or as age of an island lineage, could influence patterns of species richness if limits on total species richness imposed by the environment (i.e. ecological carrying capacities) do not exist, or if these ecological limits have not been reached (Mittelbach et al. 2007). In isolated archipelagos where most species diversity is generated in situ, island age may better predict species richness than area if speciation rates are sufficiently low such that carrying capacities have not been reached. This appears to be the case in the endemic land snail faunas of the Galapagos (Parent and Crespi 2006). Some evidence also exists for general declines in species richness over time in ecological communities (White et al. 2010), and for early post-colonization spikes in species richness, followed by decline, in both evolutionarily assembled communities (Gillespie 2004; Meyer et al. 2011), and ecological ones (Simberloff 1976). As long as carrying capacities are not reached, for diversifying clades, a positive relationship between clade age and species richness is expected even if speciation and extinction rates are exactly equivalent, and even with substantial variation in diversification rates among lineages (Rabosky 2009b, a).

Lakes, as well-circumscribed ecosystems with defined boundaries, are an aquatic equivalent to oceanic islands. However, consideration of lakes in the framework of island biogeography theory has been relatively rare. There is evidence for positive species-area relationships in freshwater fish faunas in large North American and African lakes (Barbour and Brown 1974), and in smaller temperate lake systems (Magnuson et al. 1998). Positive species-energy relationships also exist for lacustrine fish in temperate systems, but a significant positive relationship only appears at regional scales (e.g. among multiple lake), not at the scale of individual lakes (Gardezi and Gonzalez 2008). Neither species-area patterns nor species-energy patterns in fishes have been studied with regard to the influence of in situ speciation. The only suggestive evidence comes from Seehausen (2006) who showed an exponentially increasing relationship of species richness with area for individual radiating lineages of cichlid fish within African lakes. We here expand upon this dataset and examine patterns of cichlid species richness in more detail, using a dataset including 46 lakes.

Diversification of cichlid fishes within lakes across the African continent has been widespread. The East African Rift Lakes host the most species-rich radiations on the continent (Tanganyika: ~250 spp.; Malawi: 451-600 spp.; Victoria: 447-535 spp.; Genner et al. 2004), but intralacustrine diversification has also occurred in lakes from Cameroon (Schliewen et al. 1994) to the Eastern Rift, and Namibia to the Middle East (Figure 4.1). The wide geographic distribution of these lakes across the African continent means that they vary substantially in energy input from solar radiation. Fish diversity in lakes is often higher in the littoral area than in the pelagic (Vadeboncoeur et al. 2011), and thus lake perimeter might be a better metric of habitat availability than surface area for most fish groups, and perhaps especially for cichlids since the

most diverse cichlid communities are in rocky littoral habitat. However, there is a very strong correlation between lake surface area and perimeter for the lakes in the dataset included here (see Appendix 3), and thus surface area and perimeter should perform equivalently as proxies for habitat area.

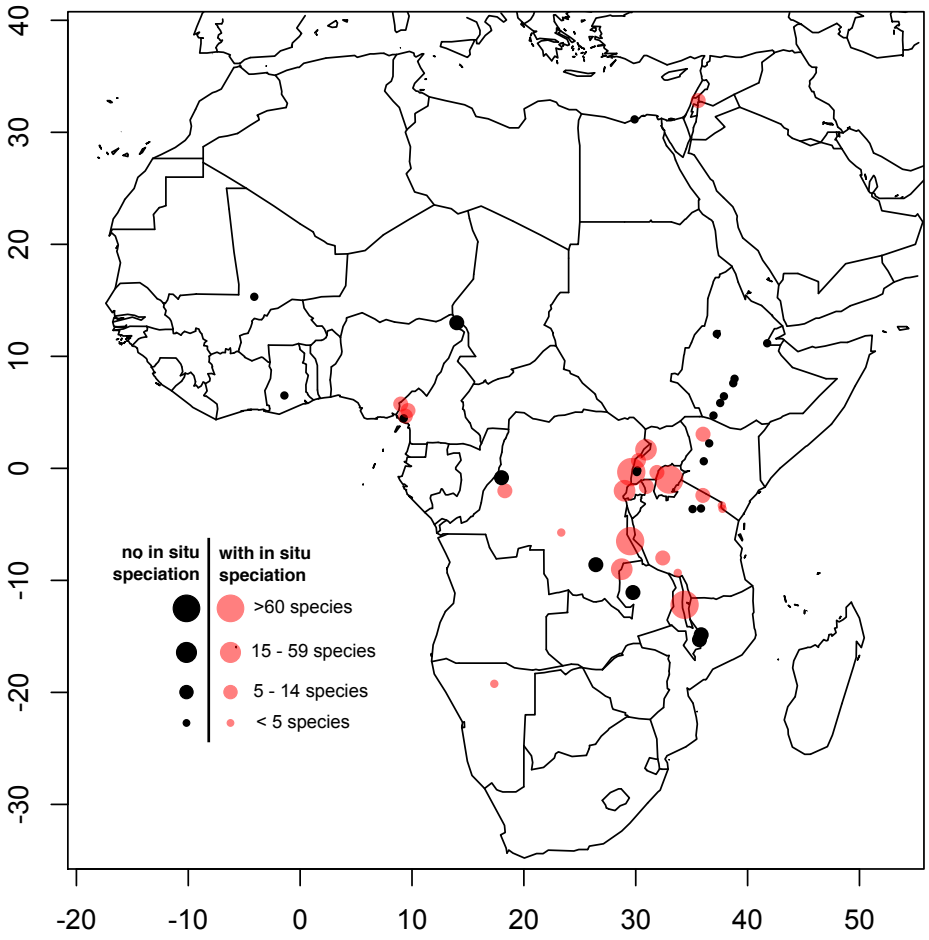


Figure 4.1. Total species richness of the cichlid faunas of 46 African lakes. Red dots indicate lakes in which in situ speciation has occurred; black dots have accumulated species richness purely via colonization.

We examine cichlid species richness patterns in African lakes at two levels (see Figure 4.2). First, in accordance with traditional island biogeography studies, we look at the total species richness of cichlids per lake. Second, we look at patterns of species richness among clades that have diversified entirely by in situ speciation within single lakes ("radiations"). The total cichlid diversity of a lake is the sum of the diversity of radiations within that lake and species entered the lake via colonization.

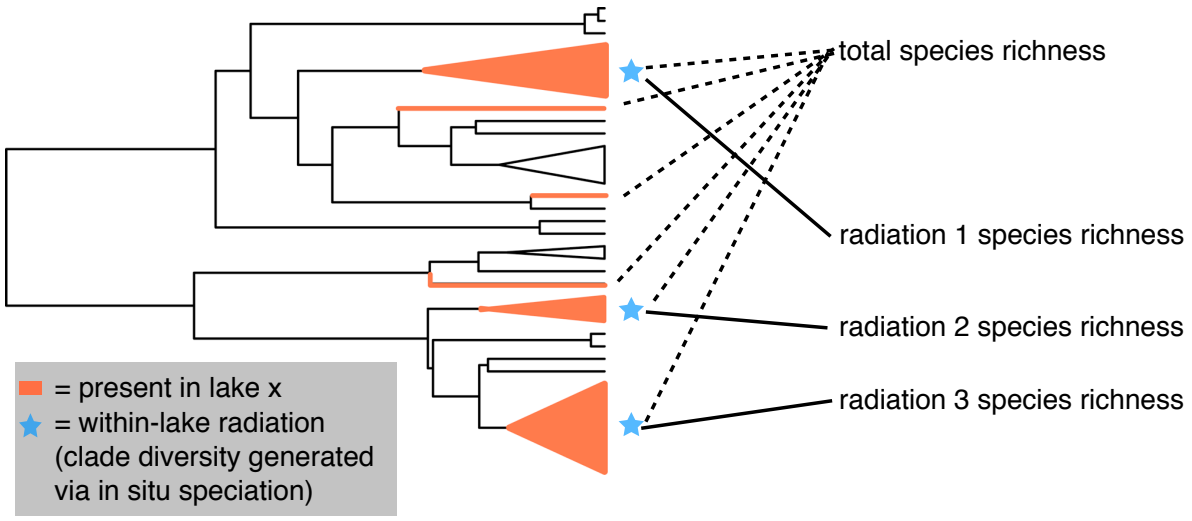


Figure 4.2. Schematic depicting the difference between total species richness and the species richness of intralacustrine radiations. Total species richness is all cichlid species present within a single lake; some species have arisen via in situ speciation, others are there via colonization. Radiations are clades that have diversified entirely by in situ speciation within single lakes. For African cichlids, it is not uncommon for multiple radiations to have occurred within a single lake.

We analyze species-area, species-energy, and species-time patterns for both radiations and for total richness of cichlids within lakes. We additionally examine the relationship between species richness and lake depth, as depth might impact total habitat area in a manner different from lake area, analogous to the influence of elevation on island species-area relationships (Ricklefs and

Lovette 1999). Using these data, we ask the following questions: 1) What is the shape of the species-area curve for cichlid faunas in African lakes, and how does in-situ speciation influence this relationship? 2) For total species richness of cichlids within lakes, what is the relationship between species richness and energy, depth, and time?; 3) For intralacustrine radiations, what is the relationship between species richness and energy, depth, and clade age?; 4) How does in situ radiation influence the total per-lake species richness patterns? Together, these questions seek to identify important environmental factors influencing species richness of cichlid fishes in African lakes, and the role of in situ diversification to shaping these patterns.

Material and Methods

We compiled information about the cichlid species present in 46 African lakes from FishBase records (Froese and Pauly 2010), and other published information (Genner et al. 2004 and references therein; Lamboj 2004). We then collated information on lake surface area and depth for these lakes (see Appendix 3, Supplementary Table 2 for values). If published records of lake surface area were unavailable, we measured surface areas from the most recent available google earth images using the software ImageJ (Rasband 1997-2011). We collated information about lake basin ages and/or time since last desiccation from a variety of sources (Appendix 4, Supplementary Table 1). We additionally calculated relative stem clade ages for lineages that diversified via within-lake speciation using ultrametric molecular phylogenies of African cichlids (for phylogenetic methods, see this volume, Chapter 3).

Our approach here is to examine patterns of species richness at two scales: 1) the total species richness of all cichlid lineages present within a lake, and 2) the richness of clades that have diversified entirely by in situ speciation within single lakes ("radiations"). Total richness is the sum of the richness of radiations, plus lineages that have entered the lake via colonization and have not diversified (see Figure 4.2). By testing the relationships of environmental variables to each of these species richness measures, we ask both how environmental variables influence the species richness of cichlid faunas of lakes, and if/how these environmental variables influence the species richness of cichlid radiations.

Testing for factors limiting species richness

It is possible that environmental variables act as factors limiting species richness. Although limiting factors influence many ecological relationships, correlation and linear regression models are not appropriate for testing relationships of this sort (Cade et al. 1999). For factors which act to limit species richness, we would expect higher predictor variable values to be positively related with the maximum observed values of species richness, but for there to be no relationship between the variable and the minimum values of species richness. We would also expect a poor fit to a linear regression model. Quantile regression can be used to test for the effects of limiting factors on the distribution of a response variable. In such a case, the upper quantiles of the relationship will produce steeper slopes that are significantly different from zero, and this slope should decrease and approach zero at lower quantiles (Cade et al. 1999; Chassot et al. 2010). In preliminary analyses, plots of radiation richnesses as a function of lake area, lake depth, and energy produced relationships of this form. We therefore used quantile regression to quantify this pattern for these variables, by calculating the 95^{th} and 5^{th} quantiles of the linear regression for

these datasets, and testing for differences in slope of these quantiles. Specifically, for a predictor variable to show evidence of being a limiting factor, for the 95th quantile of a linear regression we would expect a positive slope that is significantly different from zero, and for the 5th quantile we would expect a slope that does not differ significantly from zero.

Species-area relationships

To test for species-area relationships, and an increase in the slope of the species-area relationship for lakes above a threshold size, as in Losos and Schluter (2000), we first fitted a model with a linear relationship between species richness (S) and surface area (A), the log-transformed equivalent of the well-known power model (Lomolino 2000),

$$\log(S) = c + z * \log(A)$$

where c is the intercept of the species-area curve and z is its slope. We compared this to a model fitting species richness as a function of lake area in a two-slope regression framework. We used the formulation from Losos and Schluter (2000),

$$\log(S) = c + z1 * \log(A) + z2[(\log(A) - t) * (\log(A) \geq t)],$$

where c is the intercept, z1 is the slope of the line before the breakpoint, z2 is the slope of the line after the breakpoint, and t is the position of the breakpoint. All models were fitted using maximum likelihood estimation of nonlinear least squares parameter estimates, using the function nls() in R (R Development Core Team 2011).

To investigate whether there is a threshold size at which speciation exceeds immigration as a source of new species, we first identified lineages that had undergone at least one intralacustrine

("in situ") speciation event. We considered any endemic species co-occurring with its sister species (either a widespread species or a lake endemic itself) to have resulted from intralacustrine speciation. Single endemic species not co-occurring with a sister taxon were not considered to have arisen from in situ speciation events. We considered all fully endemic clades inhabiting a single lake to have arisen via in situ speciation. Using these criteria, we summed the total diversity of species resulting from in situ speciation. We plotted the proportion of species arising via in situ speciation to total species richness per lake (the "speciation fraction") as a function of lake size.

To investigate the influence of lake area on the species richness of intralacustrine radiations, we examined the relationship between radiation richness and lake surface area, using the quantile regression methods discussed above.

Species richness and lake depth, energy and time

To further test for environmental influences on lacustrine cichlid species richness, we examined the relationships between species richness and lake depth, energy, and time, for both total richness per lake and for the richness of intralacustrine radiations. For analyses of total per-lake species richness, we conducted linear regressions of environmental variables and species richness. If these variables influence species richness, we would expect positive correlations between environmental predictor variables and species richness. Furthermore, there could be differences in the scaling of species richness with environmental variables when there is within-lake speciation versus when communities are assembled via dispersal alone, as has been discussed for species-area patterns (e.g. Losos and Schluter 2000; Rabosky and Glor 2010;

Rosindell and Phillimore 2011). To ask whether the relationship between environmental variables and species richness differs for lakes with in situ speciation, we conducted linear regressions of lake depth, energy and time on the subset of lakes in which there has been speciation.

For the richness of radiations, we used quantile regression analyses if preliminary linear regression analyses exhibited a poor fit to the data, and plots revealed patterns characteristic of those expected for limiting factors (see discussion above). This was true for energy and for depth, but not for clade age. For clade age, we used linear regression to ask whether a relationship between clade age and species richness exists.

Combined effects of environmental factors on species richness

We looked at the combined effects of environmental variables on total species richness per lake by testing among multivariate models including all possible combinations of environmental predictor variables. We assumed linear relationships for all variables except for area, where we fitted both linear and broken regression relationships. We used AICc scores to test among models, and fit all models using the function nls() in R (R Development Core Team 2011). Because of collinearity between lake age and lake depth, we excluded lake age from all multivariate models. We tested relationships for the complete set of lakes, and for the subset of lakes with in situ speciation. Model formulations are given in Appendix 4, Supplementary Table 2.

The contribution of within-lake diversification to species richness patterns

Throughout this paper, we examine species richness of 1) lake faunas and 2) intralacustrine radiations. Here, we were interested in understanding the relationship between intralacustrine radiation and the species richness of lake faunas. Are the most species rich lake faunas dominated by single radiating clades, or do these faunas include multiple radiations? To examine the ways in which the diversity of individual lineages contributes to total cichlid diversity for lakes, we plotted the total diversity for our most species-rich lakes, and the proportion of this total diversity that each lineage (e.g. both colonizing species and intralacustrine radiations) present in the lake contributes.

Results

The total species richness of cichlid faunas varies substantially across the 46 lakes we surveyed, and half of these lakes have intralacustrine speciation (Figure 4.1).

Species-area relationships

Species richness of lake cichlid faunas shows a clear breakpoint relationship with lake area (Figure 4.3a). A two-slope regression model strongly outperforms a linear model (ΔAICc = 28). The pre-breakpoint slope of the model does not differ significantly from zero ($p = 0.411$), the breakpoint occurs at lake size 1,030 km^2, and the post-breakpoint slope is strongly positive (1.289).

Plots of speciation fraction against lake area show no relationship between the relative contribution of in-situ speciation and lake size. Speciation occurred within the smallest (0.385 km^2) and the largest (68,800 km^2) lakes in our dataset, and in lakes of all intermediate sizes (Figure 4.3b).

For radiations, as lake area increases, so does the maximum number of species observed within radiations (Figure 4.3c), and linear and two-slope regression models perform equivalently (ΔAICc 0). The slope of the linear 95th linear regression quantile differs significantly from zero ($p<0.0005$), and slope of the 5th quantile does not differ from zero ($p = 1$), suggesting that area acts as a factor limiting the species richness of radiations.

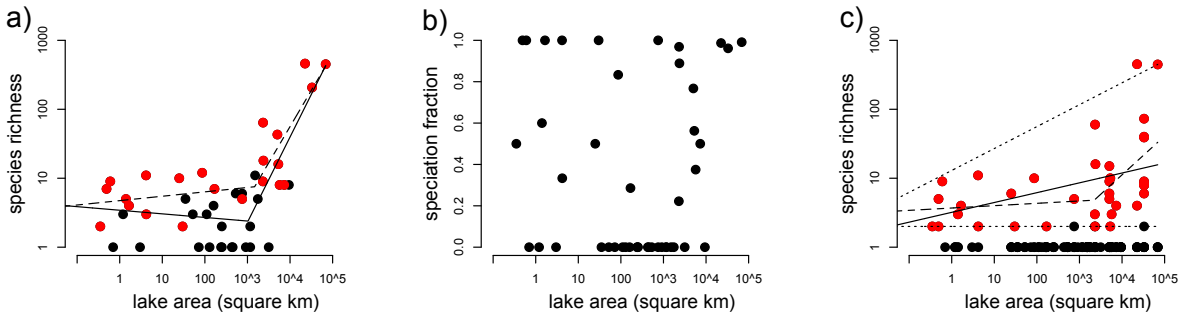


Figure 3. The species richness of cichlids in African lakes is best explained by a two-slope breakpoint regression model, but this breakpoint model does not correspond with a size-threshold for in-situ speciation. **a)** For lakes above 1,030 km^2 there is a strong positive relationship between lake surface area and species richness (slope = 1.289). For lakes below this threshold size, there is no relationship between species richness and lake area. Solid line corresponds to the two-slope regression model for the whole dataset; dashed line is the two-slope model for lakes with in situ speciation only (red dots). Neither pre-breakpoint slope differs significantly from zero. **b)** The proportion of total species in a lake fauna that has arisen via in-situ speciation ("speciation fraction"), as a function of lake area. Speciation occurs within both the smallest and the largest lakes in the dataset. **c)** Species richness of radiations in African lakes is limited by lake area. Solid and dashed lines represent the best fit linear and two-slope regression models for radiations only (red dots). Black dots represent colonizing lineages that have not diversified within the lake. Dotted lines are the 5th and 95th quantiles of the linear model.

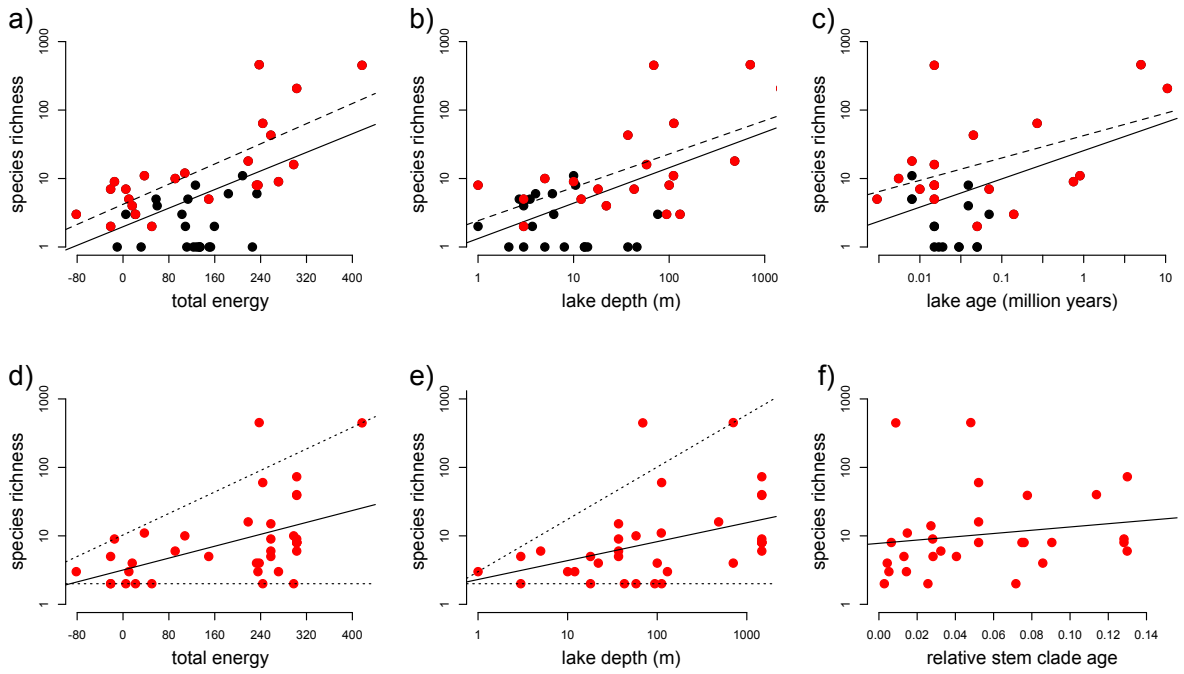


Figure 4. There are significant positive correlations between species richness of cichlids in African lakes and a) total energy ($r^2 = 0.296$), b) lake depth ($r^2 = 0.330$), and c) lake age ($r^2 = 0.559$). For cichlid radiations (panels d-f), species richness shows evidence of limitation by both d) total energy and e) lake depth. For panels a-c, solid lines are the result of linear regression for the entire dataset; dashed lines are those for the subset of lakes wherein there is speciation (red dots). The correlation between energy and species richness substantially improves when only including lakes with in-situ speciation (red dots; $r^2 = 0.559$). For panels d-f, solid lines are the results of linear regression; dotted are the 5th and 95th quantiles of these datasets. f) For cichlid radiations, there is no relationship between clade age and species richness ($r^2 = 0.022$, $p = 0.429$), nor is there evidence for clade age-based constraint on lineage species richness.

Species richness and lake depth, energy and time

In analyses of the relationships between total per-lake species richness and energy, lake depth, and age, we find significant positive linear relationships (Figure 4.4a-c). There are significant correlations between species richness and energy (E_{tot}; r2 = 0.296), lake depth ($r^2 = 0.330$) and lake age ($r^2 = 0.295$). When we exclude lakes in which there is not in situ speciation, these

relationships remain significant and positive, and for energy the correlation substantially improves ($r^2 = 0.559$).

For the species richness of radiations, plots reveal a pattern of limitation by energy and depth: these variables predict the maximum observed species richness of these clades (Figure 4.4d, 4.4e). Quantile regression analyses provides support for both energy and depth acting as limiting factors; for both of these variables, the slope of the 95th quantile is positive and differs significantly from zero ($p < 0.05$ for both variables), but the slope of the 95th quantile does not differ from zero ($p = 1$ for both variables). In contrast, there is no relationship between relative clade age of radiations and species richness, and no evidence for a limitation of species richness by clade age (Figure 4.4f)**.**

Combined effects of environmental factors on species richness

Multivariate models revealed that combinations of environmental predictor variables, not single predictors, best explained variation in species richness for the cichlid faunas of African lakes (Table 4.1). For the complete dataset, a model including the two-slope species area relationship plus lake depth performed best. For the subset of lakes with in situ speciation, the best-performing model was one including linear area, depth and total energy terms.

Note that we did not test multivariate models for the species richness of radiations. Because depth and energy show evidence of acting as limiting factors, and not predictors, of species richness, they violate the assumptions of linear models and therefore using them in multivariate linear models would not be appropriate. Multivariate analytical methods for limiting factors have, to our knowledge, not been developed.

Table 4.1. Multiple regression models reveal that models incorporating multiple environmental variables outperform those with lake surface area alone. For the total dataset, the best model is one with a two-slope area term and linear depth term. For lakes with in situ speciation, a model including linear relationships with area, depth and total energy performs best. Other than where indicated as "two-slope", all terms are linear.

Full Dataset, total per-lake richness, n=46			**Lakes with in-situ speciation,** n=23	
	AICc	ΔAICc	AICc	ΔAICc
Area	88.50	37.52	36.40	13.17
Depth	72.87	21.89	40.01	16.78
Total Energy	73.31	22.33	35.81	12.58
two-slope Area	60.08	9.10	27.38	4.15
two-slope Area + Depth	**50.98**	**0.00**	**23.23**	**0.00**
two-slope Area + Energy	63.29	12.31	30.47	7.24
two-slope Area + Energy + Depth	59.43	8.45	26.54	3.31
Depth + Energy	56.88	5.90	28.12	4.89

The contribution of within-lake diversification to species richness patterns

We examined how the diversity of radiating clades contributes to the diversity of lake faunas by examining the most species-rich lakes in our dataset. All of these lakes were above 1,030 km^2, the threshold point in the broken regression that best fit the species-area pattern for lake cichlid faunas. In these lakes, high richness is achieved both by high levels of diversification within single radiations (e.g. Lakes Malawi and Victoria) and by accumulation of more moderate diversity within several radiations (e.g. Lakes Tanganyika and Mweru) (Figure 4.5). Therefore, high species richness of lake faunas is achieved by a both extensive radiation of a single colonizing species, and by radiation in multiple colonizers.

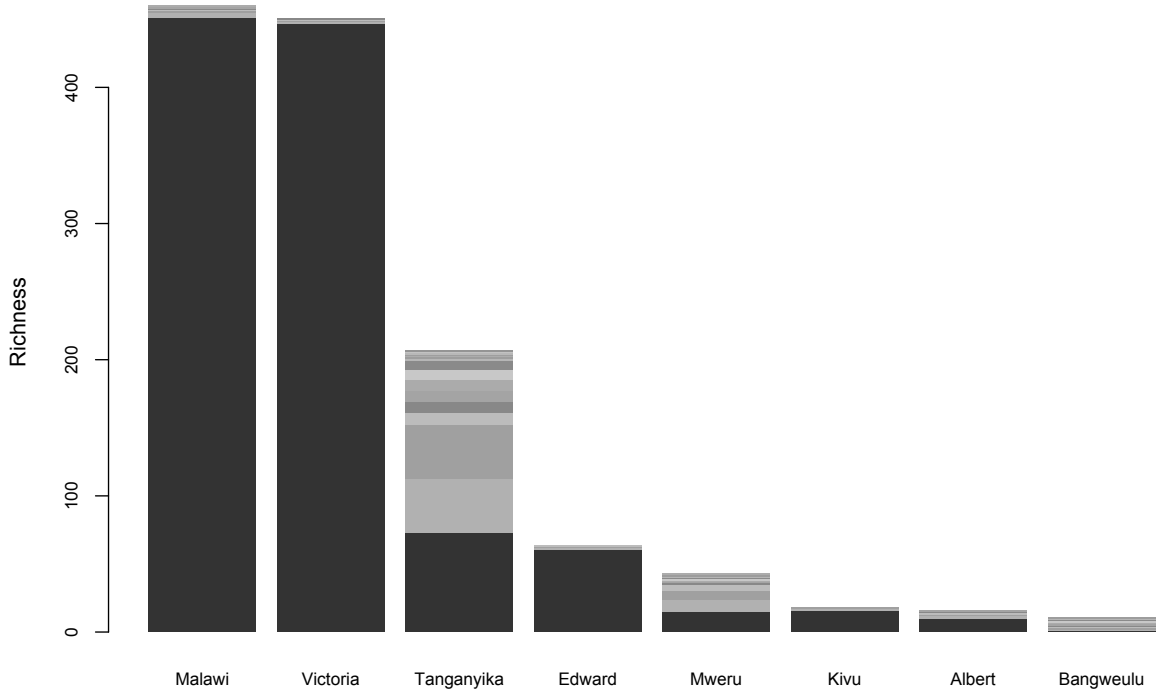


Figure 4.5. Species richness in the eight most species-rich lakes in the dataset reveal that lakes accumulate richness both through dramatic diversification of single lineages (e.g. Lakes Malawi, Victoria and Edward) or via accumulation of diversity in many lineages (e.g. Lakes Tanganyika, Mweru). Each bar corresponds to a lake; different colors within each bar represent the diversity of the lineages present within that lake. The lakes shown here are the most species-rich in the dataset; all are above 1,030 km^2 in surface area.

Discussion

Our results show that variation in cichlid species richness in African lakes is correlated with measures of lake area, lake depth, and energy availability, particularly in lakes that harbor diversifying lineages. This suggests that metrics of lake size and energy are predictors of diversity for the evolutionarily assembled cichlid faunas of African lakes. These same environmental variables limit species richness of within-lake radiations. Although adaptive radiations are commonly assumed to follow niche-filling diversification processes, rarely have

the environmental variables that functionally limit species richness been considered or identified. Combining the results of analyses at the radiation richness level and at the total per-lake species richness level, we show that high species richness of cichlids within lakes is achieved by both high richness of a single predominant radiating group in some cases, and by intermediate richness of multiple radiations in other cases. These results are consistent with the existence of ecological carrying capacities which render the total diversity of cichlids predictable within these lakes. However, these carrying capacities are achieved by lineage-specific diversification outcomes, which produce marked variation in faunal composition among lakes.

Species-area patterns

Analyses of species-area relationships for total cichlid richness within lakes reveal a two-slope breakpoint relationship with lake area. Below the threshold size of ~1,030 km^2, there is no relationship between cichlid species richness and lake area, but after this threshold point the species-area relationship takes on a strong positive slope (Figure 4.3a). However, unlike previous work that has theorized breakpoint models in species-area relationships in evolutionary systems (Losos and Schluter 2000; Losos and Parent 2010), this threshold does not correspond with the area at which within-lake speciation begins. For cichlids, there is no relationship between lake size and the proportion of taxa that have evolved within the lake (Figure 4.3b). African cichlids speciate within the smallest lakes in our dataset (<1 km^2), as well as in the largest (Lake Victoria, 68,800 km^2). When we consider only lakes with in situ speciation, a breakpoint model still strongly outperforms a linear regression model, and model parameters do not substantially differ from those inferred for the complete dataset.

An additional contrast in the breakpoint species-area relationship in cichlids versus those discussed in other evolutionary systems is that we find no relationship between species richness and lake area prior to the breakpoint. In pattern, this is analogous to the well-documented "small island effect" in island biogeography where islands below a certain threshold size do not show a predictable increase in species richness with area (Lomolino and Weiser 2001). Potential hypotheses to explain the small island effect include explanations related to biases in the sampling of habitat diversity below a size threshold (Sfenthourakis and Triantis 2009), and unpredictable extinction rates due demographic stochasticity in small populations (MacArthur and Wilson 1963). However, because the threshold point in our analysis lies at a fairly large lake size, both of these explanations seem unlikely for cichlids in these lakes. Additionally, a breakpoint relationship cannot result from a lake size-derived increase in the number of colonizing lineages alone (see Appendix 4).

Although the onset of in situ speciation cannot explain the breakpoint species-area relationship for cichlids, differences in modes of speciation with area could contribute to this pattern. Mechanisms of cichlid speciation can operate at very small spatial scales (e.g. Seehausen et al. 2008) and cichlids are known to speciate in very small lakes (Schliewen et al. 1994; Barluenga et al. 2006); this study confirms that in situ speciation is not inhibited by lake size within the size range that analyzed here (see also Chapter 3, this volume). However, cichlids also speciate geographically, and a number of studies have shown that rocky habitat cichlids can have extremely fine-scale population genetic structure which may contribute to this propensity (Arnegard et al. 1999; Markert et al. 1999; Danley et al. 2000; Wagner and McCune 2009; Losos and Parent 2010). If rates of allopatric speciation increase with area, and non-geographic

speciation happens consistently at all lake sizes, the interaction between the addition of species to the system via these different modes of speciation could create a non-linear species-area relationship.

Note also that for cichlids, it is likely that lake perimeter is actually a better predictor of the probability of allopatric speciation than lake area, because the vast majority of fishes use littoral, not pelagic habitat (Vadeboncoeur et al. 2011). However, because there is very strong correlation between perimeter and area for lakes in our dataset (see Appendix 4, Supplementary Figure 3), these two measures are effectively interchangeable. This also means that the relationship between species richness and area that we observe could be driven by mechanisms operating via the influence of perimeter, and not area per se.

The distribution of species richness of intralacustrine radiations provides evidence that lake area limits species richness (Figure 4.3c). These clade-specific richness outcomes are likely influenced by the effect of lineage-specific traits on diversification. Traits are well known to influence diversification rates (e.g. Rabosky and McCune 2010), and traits like the intensity of sexual selection (Kraaijeveld et al. 2010), sexual conflict (e.g. Arnqvist et al. 2000), or degree of ecological specialization (e.g. Farrell 1998), might influence cichlid diversification patterns. Chapter 3, this volume, shows that sexual dichromatism, a proxy for intensity of sexual selection, is correlated with instances of cichlid radiation in these same African lakes; this trait may also influence diversification rates and species richness. Variation in species richness among diversifying lineages could also result from historical contingencies (e.g. Seehausen 2007), including lineage priority effects, where early-colonizing lineages, with a head-start on diversification, would be capable of producing more species than later-colonizing lineages. Such

priority effects, combined with differences in diversification rates, would produce marked variation in the diversity of lineages that comprise lake faunas, as we observe.

Evidence from a number of systems suggests that species richness increases with increasing area at a faster rate in communities assembled by evolution (e.g. within which there is speciation) than those assembled purely by dispersal (Rosenzweig 1995; Triantis et al. 2008). In our dataset, lakes with in situ diversification consistently have higher species richness than do lakes without (Figure 4.3a; red dots versus black dots, respectively)**.** This pattern is consistent with a faster rate of species richness increase with area in evolution-assembled versus dispersal-assembled cichlid communities. However, it could also be that lakes without in situ speciation differ from lakes with speciation in a way which compromises the comparability of these groups; it is possible that the absence of intralacustrine speciation is reflective of non-equilibrium conditions, and thus a difference between these communities and those in equilibrium would be expected.

Species-energy patterns

Total energy per lake (e.g. net solar radiation multiplied by lake area) is significantly positively correlated with total cichlid species richness in lakes, and this relationship is linear (Figure 4.4a). This correlation improves when we look only at lakes within which there has been speciation. As discussed above for area, where we see a parallel pattern, this might suggest that evolutionarily-assembled communities can contain more species per given "energy-area" than can dispersal assembled communities. This would suggest that species originating via in situ speciation were able to more effectively exploit resources, perhaps via tighter niche packing. However, this could

also reflect nonequilibrium conditions in lakes without in situ speciation, compared to lakes with in situ speciation.

Differences in energy availability almost certainly influence diversity through their effect on primary productivity in these aquatic systems (Wetzel 2001). The relationship between diversity and productivity in aquatic systems is well-established as being highly dependent upon spatial scale, as it is in terrestrial systems (Mittelbach et al. 2001); at local scales it is hump-shaped, but at regional scales it becomes linear, due to interactions between local and regional processes influencing community structure and diversity (Chase and Leibold 2002). Thus, a linear relationship, given the large-grained scale of this study, is expected. However, prior evidence for correlation between fish species richness and energy at the lake level is lacking, although positive linear species-energy relationships emerge at regional scales in one temperate lake system where this pattern was extensively studied (Gardezi and Gonzalez 2008). Ours is the first evidence, to our knowledge, that energy predicts species richness in an island system in which evolution contributes substantially to total species diversity.

For intralacustrine radiations, energy acts as a factor limiting species richness (Figure 4.4d). Because the diversity of these lineages has accumulated via in situ speciation, this is evidence for linkage between diversification and energy, as is often hypothesized (Currie et al. 2004; Evans et al. 2005) but rarely tested. This linkage could be through direct influences on diversification rate, via increases in mutation rate and/or generation time (Evans and Gaston 2005), or via decreased extinction rates and/or increased speciation rates mediated by larger per area population sizes with increased energy availability (Evans et al. 2005). However, lack of complete species-level phylogenies constrains our ability to test directly for linkages between diversification rates and

energy availability in this system. Furthermore, if carrying capacities exist, meaningful tests of diversification rates are impossible using phylogenies of extant taxa, unless equilibrium carrying capacities have not yet be reached (Rabosky 2009a).

Other proposed mechanisms linking energy availability and species richness are more explicitly ecological mechanisms. However, these mechanisms could also influence speciation and evolutionary dynamics in these systems. Ecological mechanisms include the idea that higher energy systems contain species with reduced niche breadth and/or more specialized niche position, and the idea that more trophic levels are exploitable in higher energy systems (Evans et al. 2005). These mechanisms would be readily testable if detailed information about trophic ecology of cichlid communities from both high and low-energy systems were available. Cichlids are well-known for their often highly specialized trophic ecology (Fryer and Iles 1972; Liem 1973; Genner et al. 1999), but detailed information about trophic community ecology is lacking, especially for smaller, lower-energy lakes.

Depth and Time

Both lake depth and lake age are significantly positively correlated with total per-lake cichlid species richness. However, depth and lake age are also collinear, especially for deep lakes (r^2 = 0.43; see Appendix 4, Supplementary Figure 4), as deeper lakes are resistant to desiccation and persist longer than shallow lakes. Depth might also mediate increased species richness via increasing available habitat, perhaps in an analogous manner to elevation's use as a metric of habitat diversity on islands (Ricklefs and Lovette 1999). As diversity in fishes can be depth-structured, depth is likely an important dimension of habitat (e.g. Ingram and Shurin 2009;

Vonlanthen et al. 2009). Alternatively, the correlation of lake age and depth with species richness might suggest that deeper, older lakes simply have more time to accumulate diversity, an interpretation that would imply that these lakes are not at equilibrium diversity levels. These two explanations are not mutually exclusive, and both processes could contribute to observed species richness patterns.

For intralacustrine radiations we observe no relationship between relative clade age and species richness. Even if speciation and extinction rates are equal, a positive relationship between clade age and species richness is expected, yet many studies report no such relationship (Rabosky 2009b). One explanation for this pattern is that there exist environmental constraints on total diversification, thus rendering speciation and extinction rates diversity-dependent. Diversity-dependent diversification is expected if speciation or extinction rates depend on resource availability (Walker and Valentine 1984). This is a commonly discussed pattern, and some definitions of adaptive radiation even include slowdown in diversification rates over time as an expected by-product of niche-filling processes (Schluter 2000). The little-discussed corollary to this idea is that the species richness of adaptive radiations should thus be reflective of ecological carrying capacities, if these equilibriums have been reached. We view this as a likely explanation for the lack of clade age-species richness relationship for cichlids in African lakes.

Combined effects of environmental factors on species richness

Testing among multivariate models incorporating all combinations of environmental predictor variables revealed that the best fitting models were those incorporating multiple predictors. For the complete dataset, the best fitting model was one incorporating the two-slope broken

regression of lake area, combined with lake depth. Because surface area and depth together are a metric for total habitat area, this result suggests that habitat area predicts the diversity of lake cichlid faunas, at least in lakes above the threshold area size.

For the subset of lakes with in situ speciation, a model incorporating area, depth and total energy performed best. This suggests that the most important factors predicting species richness in faunas with in situ speciation may be slightly different than for the total dataset. In particular, total energy is included as a predictor in the model for lakes with in situ speciation, suggesting that energy and habitat area together best predict the diversity of lakes with diversification. This is consistent with the finding of Chapter 3, this volume that energy is an important factor in predicting cichlid radiations in lakes, along with lake depth, lake area, and sexual dichromatism.

Summary and Conclusions

African cichlids exhibit a breakpoint species-area relationship, but the breakpoint cannot be explained by the onset of in situ speciation, the mechanism attributed to this pattern in other cases in which it has been observed (Losos and Schluter 2000; Losos and Parent 2010). In cichlids, this may result from an interaction between geographic and nongeographic speciation mechanisms, if frequency of geographic speciation scales with lake area or perimeter.

Total cichlid species richness is significantly linearly related to energy, lake depth and lake age. This is the first evidence that energy is related to, and predicts, species richness in an "island" system in which evolution contributes substantially to total species diversity. Furthermore, multivariate models indicate that multiple predictor variables, including lake area, lake depth and total energy, best explain species richness of cichlid faunas within African lakes. We suggest that

metrics of lake size and energy function as predictors of carrying capacity of evolutionarily-assembled cichlid faunas in African lakes.

The species richness of intralacustrine radiations is limited by lake size (e.g. area and depth) and energy. Environmental variables limit the maximum richness of these clades, but do not predict the extent to which this maximum richness is reached. This indicates that other factors, such as lineage-specific traits, or historical contingencies, also have a strong influence on the species richness of African cichlid radiations. Although niche-filling processes are often invoked as fundamental to the theory of adaptive radiation, rarely have the environmental variables that functionally limit species richness been directly investigated. We suggest that lake size and energy function as ecological limits to cichlid adaptive radiation.

Lineage-specific diversification outcomes produce substantial variation in the way in which total cichlid species richness per lake is distributed among lineages. In some lakes total richness is almost entirely the result of dramatic diversification of single lineages (e.g. Malawi, Victoria and Edward), whereas in other lakes species richness is partitioned more evening among several lineages (e.g. Tanganyika, Mweru). Although differences in diversification rates cannot explain differences in the total species richness of cichlids in African lakes if carrying capacities limit maximum richness (e.g. Rabosky 2009a), differences in diversification rate could strongly influence how total species richness is partitioned among the diversifying lineages of a lake.

Acknowledgements

This chapter was written in collaboration with Ole Seehausen and Luke Harmon. Thanks to Amy McCune, Irby Lovette, Rick Harrison, Ellinor Michel, and Blake Matthews for comments on this manuscript, and to Pete McIntyre, Dan Rabosky, Ben Dalziel for helpful discussions.

REFERENCES

Arnegard, M. E., J. A. Markert, P. D. Danley, J. R. Stauffer, A. J. Ambali, and T. D. Kocher. 1999. Population structure and colour variation of the cichlid fish *Labeotropheus fulleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. Proc. R. Soc. B 266:119-130.

Arnqvist, G., M. Edvardsson, U. Friberg, and T. Nilsson. 2000. Sexual conflict promotes speciation in insects. Proc. Natl. Acad. Sci. USA 97:10460-10464.

Barbour, C. D., and J. H. Brown. 1974. Fish species diversity in lakes. Am. Nat. 108:473-489.

Barluenga, M., K. N. Stolting, W. Salzburger, M. Mushick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439:719-723.

Cade, B. S., J. W. Terrell, and R. L. Schroeder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology 80:311-323.

Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity-biodiversity relationship. Nature 416:427-430.

Chassot, E., S. Bonhommeau, N. K. Dulvy, F. Melin, R. Watson, D. Gascuel, and O. Le Pape. 2010. Global marine primary production constrains fisheries catches. Ecol. Lett. 13:495-505.

Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J. F. Guegan, B. A. Hawkins, D. M. Kaufman, J. T. Kerr, T. Oberdorff, E. O'Brien, and J. R. G. Turner. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecol. Lett. 7:1121-1134.

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

Evans, K. L., and K. J. Gaston. 2005. Can the evolutionary-rates hypothesis explain species-energy relationships? Funct. Ecol. 19:899-915.

Evans, K. L., P. H. Warren, and K. J. Gaston. 2005. Species-energy relationships at the macroecological scale: a review of the mechanisms. Biological Reviews 80:1-25.

Farrell, B. D. 1998. "Inordinate fondness" explained: Why are there so many beetles? Science 281:555-559.

Froese, R., and D. Pauly. 2010. FishBase.

Fryer, G., and T. D. Iles. 1972. The cichlid fishes of the great lakes of Africa: their biology and evolution. Tropical Fish Hobbyist Publications Inc., Edinburgh.

Gardezi, T., and A. Gonzalez. 2008. Scale dependence of species-energy relationships: Evidence from fishes in thousands of lakes. Am. Nat. 171:800-815.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405:220-227.

Genner, M. J., O. Seehausen, D. F. R. Cleary, M. E. Knight, E. Michel, and G. F. Turner. 2004. How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? J. Biogeogr. 31:93-102.

Genner, M. J., G. F. Turner, and S. J. Hawkins. 1999. Foraging of rocky habitat cichlid fishes in Lake Malawi: coexistence through niche partitioning? Oecologia 121:283-292.

Gillespie, R. 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science 303:356-359.

Hurlbert, A. H., and W. Jetz. 2010. More than "More Individuals": The Nonequivalence of Area and Energy in the Scaling of Species Richness. Am. Nat. 176:E50-E65.

Ingram, T., and J. B. Shurin. 2009. Trait-based assembly and phylogenetic structure in northeast Pacific rockfish assemblages. Ecology 90:2444-2453.

Kraaijeveld, K., F. J. L. Kraaijeveld-Smit, and M. E. Maan. 2010. Sexual selection and speciation: the comparative evidence revisited. Biol. Rev. Camb. Philos. Soc.

Lamboj, A. 2004. The cichlid fishes of western Africa. The cichlid fishes of western Africa.: 1-255.

Liem, K. F. 1973. Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. Systematic Zoology 22:425-441.

Lomolino, M. V. 2000. Ecology's most general, yet protean pattern: the species-area relationship. J. Biogeogr. 27:17-26.

Lomolino, M. V., and M. D. Weiser. 2001. Towards a more general species-area relationship: diversity on all islands, great and small. J. Biogeogr. 28:431-445.

Losos, J. B., and C. E. Parent. 2010. The speciation-area relationship. The theory of island biogeography revisited.:415-438.

Losos, J. B., and R. E. Ricklefs. 2010. The theory of island biogeography revisited. The theory of island biogeography revisited.:i-xvi, 1-476.

Losos, J. B., and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. Nature 408:847-850.

MacArthur, R. H., and E. O. Wilson. 1963. Equilibrium-theory of insular zoogeography. Evolution 17:373-&.

Magnuson, J. J., W. M. Tonn, A. Banerjee, J. Toivonen, O. Sanchez, and M. Rask. 1998. Isolation vs. extinction in the assembly of fishes in small northern lakes. Ecology 79:2941-2956.

Markert, J. A., M. E. Arnegard, P. D. Danley, and T. D. Kocher. 1999. Biogeography and population genetics of the Lake Malawi cichlid Melanochromis auratus: habitat transience, philopatry and speciation. Mol. Ecol. 8:1013-1026.

Meyer, J. R., S. E. Schoustra, J. Lachapelle, and R. Kassen. 2011. Overshooting dynamics in a model adaptive radiation. Proc. R. Soc. B 278:392-398.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, A. H. Hurlbert, N. Knowlton, H. A. Lessios, C. M. McCain, A. R. McCune, L. A. McDade, M. A. McPeek, T. J. Near, T. D. Price, R. E. Ricklefs, K. Roy, D. F. Sax, D. Schluter, J. M. Sobel, and M. Turelli. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecol. Lett. 10:315-331.

Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig, S. I. Dodson, and L. Gough. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381-2396.

Parent, C. E., and B. J. Crespi. 2006. Sequential colonization and diversification of Galapagos endemic land snail genus Bulimulus (Gastropoda, Stylommatophora). Evolution 60:2311-2328.

Rabosky, D. L. 2009a. Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. Ecology Letters 12:735-743.

Rabosky, D. L. 2009b. Ecological Limits on Clade Diversification in Higher Taxa. American Naturalist 173:662-674.

Rabosky, D. L., and R. E. Glor. 2010. Equilibrium speciation dynamics in a model adaptive radiation of island lizards. Proc. Natl. Acad. Sci. USA 107:22178-22183.

Rabosky, D. L., and A. R. McCune. 2010. Reinventing species selection with molecular phylogenies. Trends Ecol. Evol. 25:68-74.

Rasband, W. S. 1997-2011. ImageJ. U.S. National Institutes of Health, Bethesda, MD, USA.

Ricklefs, R. E., and I. J. Lovette. 1999. The roles of island area per se and habitat diversity in the species-area relationships of four Lesser Antillean faunal groups. J. Anim. Ecol. 68:1142-1160.

Rosenzweig, M. L. 1992. Species-diversity gradients – we know more and less than we thought. J. Mammal. 73:715-730.

Rosenzweig, M. L. 1995. Species diversity in space and time. Species diversity in space and time.:i-xxi, 1-436.

Rosenzweig, M. L. 2001. Loss of speciation rate will impoverish future diversity. Proc. Natl. Acad. Sci. USA 98:5404-5410.

Rosindell, J., and A. B. Phillimore. 2011. A unified model of island biogeography sheds light on the zone of radiation. Ecol. Lett.

Schliewen, U. K., D. Tautz, and S. Paabo. 1994. Sympatric Speciation Suggested By Monophyly Of Crater Lake Cichlids. Nature 368:629-632.

Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, Oxford.

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

Seehausen, O. 2007. Evolution and ecological theory - Chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. Heredity 99:361-363.

Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V. Schneider, M. E. Maan, H. Tachida, H. Imai, and N. Okada. 2008. Speciation through sensory drive in cichlid fish. Nature 455:620-U623.

Sfenthourakis, S., and K. A. Triantis. 2009. Habitat diversity, ecological requirements of species and the Small Island Effect. Divers. Distrib. 15:131-140.

Simberloff, D. 1976. Species turnover and equilibrium island biogeography. Science 194:572-578.

Storch, D., K. L. Evans, and K. J. Gaston. 2005. The species-area-energy relationship. Ecol. Lett. 8:487-492.

Team, R. D. C. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Triantis, K. A., M. Mylonas, and R. J. Whittaker. 2008. Evolutionary species-area curves as revealed by single-island endemics: insights for the inter-provincial species-area relationship. Ecography 31:401-407.

Vadeboncoeur, Y., P. B. McIntyre, and M. J. Vander Zanden. 2011. Borders of Biodiversity: Life at the Edge of the World's Large Lakes. Bioscience 61:526-537.

Vonlanthen, P., D. Roy, A. G. Hudson, C. R. Largiader, D. Bittner, and O. Seehausen. 2009. Divergence along a steep ecological gradient in lake whitefish (Coregonus sp.). J. Evol. Biol. 22:498-514.

Wagner, C. E., and A. R. McCune. 2009. Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. Evolution 63:1312-1326.

Walker, T. D., and J. W. Valentine. 1984. Equilibrium models of evolutionary species diversity and the number of empty niches. Am. Nat. 124:887-899.

Wetzel, R. G. 2001. Limnology: Lake and River Ecosystems, Third Edition. Academic Press.

White, E. P., S. K. M. Ernest, P. B. Adler, A. H. Hurlbert, and S. K. Lyons. 2010. Integrating spatial and temporal approaches to understanding species richness. Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci. 365:3633-3643.

Wright, D. H. 1983. Species-energy theory – an extension of species-area theory. Oikos 41:496-506.

APPENDIX 1

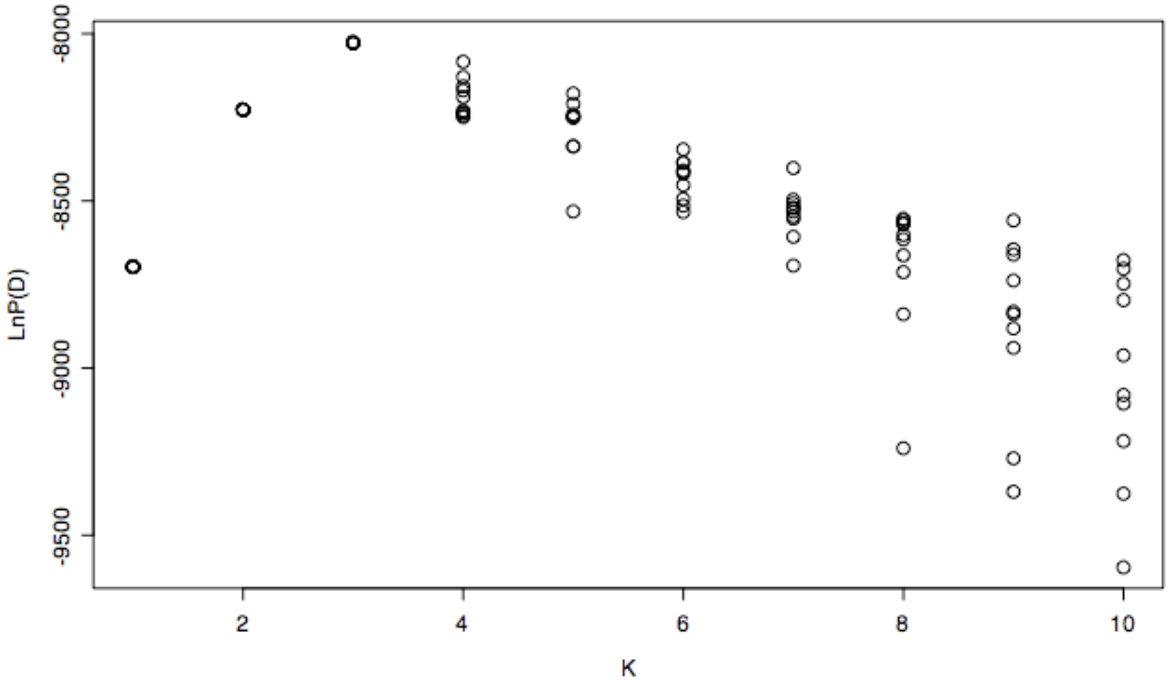


Figure S1. Log probability of the data (LnP(D)) for 10 STRUCTURE runs at each K value, for K 1 through 10 for *P.* sp. "kazumbe".

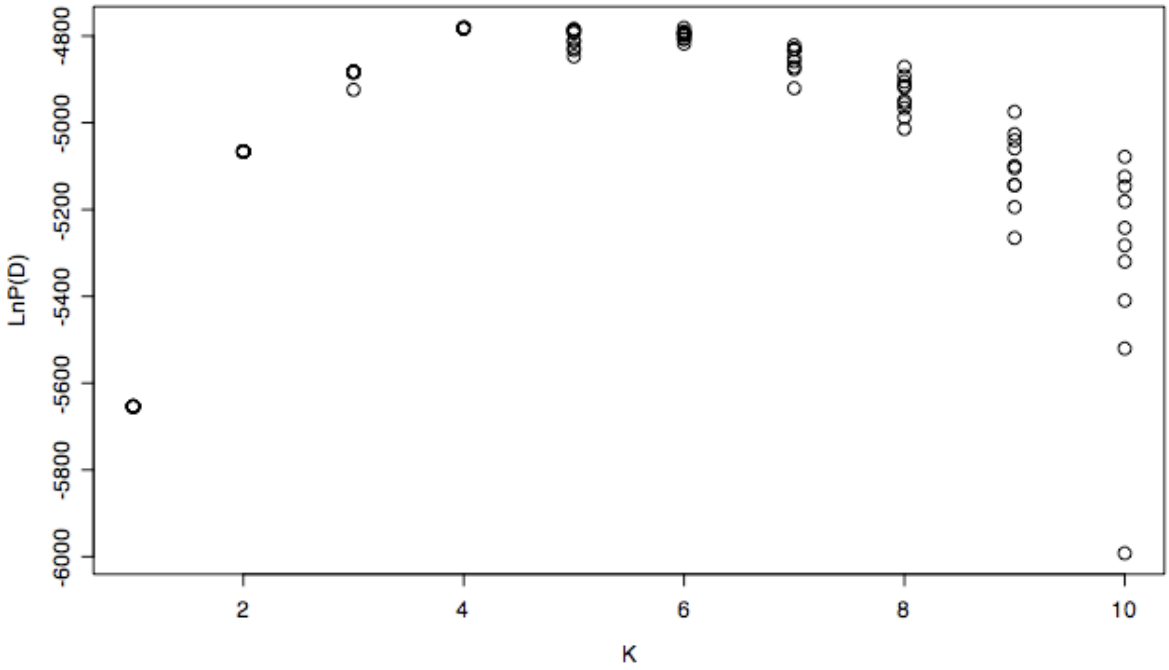


Figure S2. Log probability of the data (LnP(D)) for 10 STRUCTURE runs at eack K value, for K 1 through 10 for *P.* sp. "moshi".

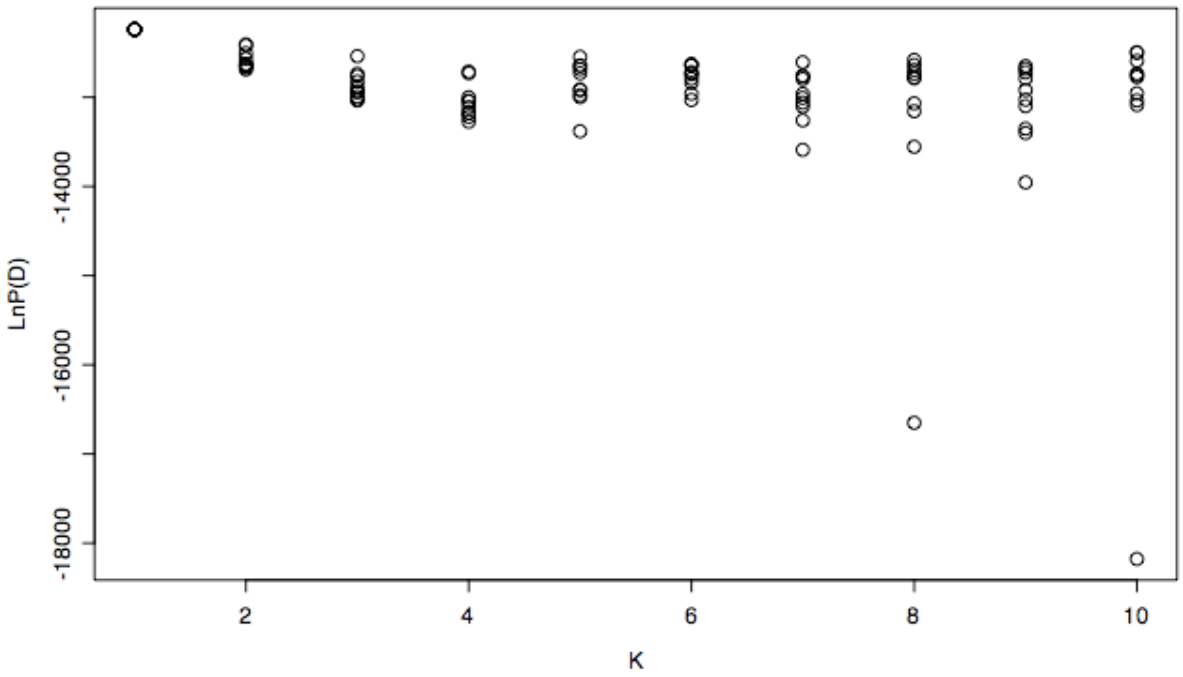


Figure S3. Log probability of the data (LnP(D)) for 10 STRUCTURE runs at eack K value, for K 1 through 10 for *S. diagramma*.

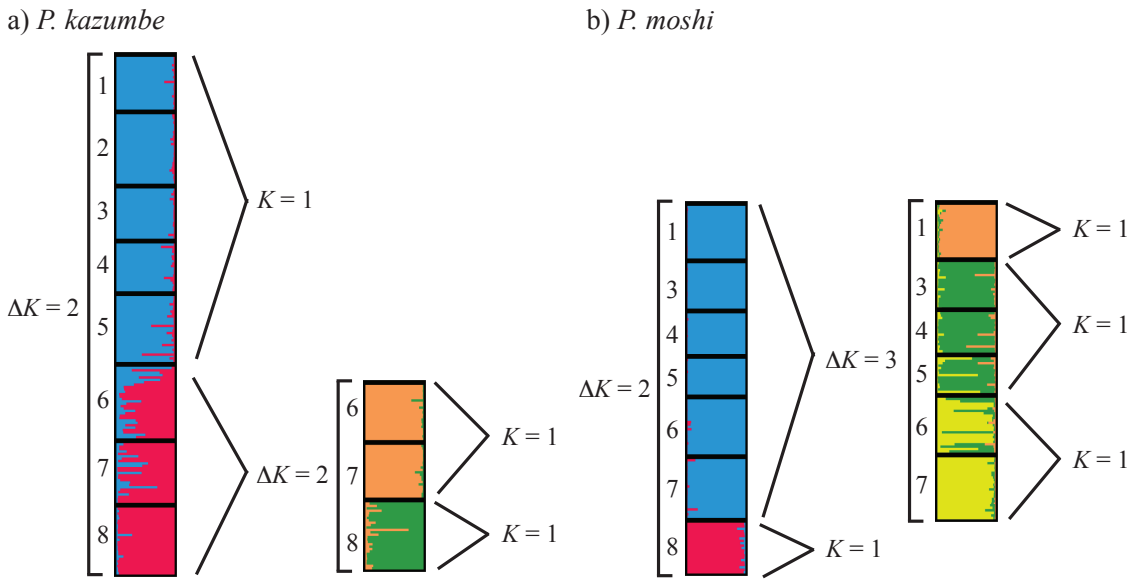


Figure S4. Results of fully hierarchical Δ*K* STRUCTURE analyses for a) *P.* sp. "kazumbe" and b) *P.* sp. "moshi". At each round, datasets were divided into *K* subsets, where *K* is the number of genetic groups supported by Δ*K* analyses (Evanno et al 2005) of the dataset from the previous round. Individuals were assigned into subsets if their assignment probabilities were 0.6 or higher for a given group. Assignment probabilities used for subsetting were the consensus of 10 runs at the *K* indicated by Δ*K* analyses, generated using CLUMPP (Jakobsson and Rosenberg 2007).

Table S1. All individuals used in this study have been deposited in the Cornell Museum of Vertebrates (CUMV) as voucher specimens. Identification numbers for all individuals, and CUMV numbers linked to these specimens and tissue vouchers, are given below.

Species	Individual ID	CUMV Number	Species	Individual ID	CUMV Number	Species	Individual ID	CUMV Number
Petrochromis sp. "*kazumbe*"	CEW07F.171	93663	*Petrochromis* sp. "*moshi*"	CEW07F.174	93669	*Simochromis diagramma*	CEW07F.205	93667
	CEW07F.173	93663		CEW07F.177	93669		CEW07F.206	93667
	CEW07F.175	93663		CEW07F.181	93669		CEW07F.207	93667
	CEW07F.180	93663		CEW07F.184	93669		CEW07F.208	93667
	CEW07F.183	93663		CEW07F.186	93669		CEW07F.213	93667
	CEW07F.185	93663		CEW07F.195	93669		CEW07F.217	93667
	CEW07F.187	93663		CEW07F.196	93666		CEW07F.219	93667
	CEW07F.190	93663		CEW07F.197	93666		CEW07F.220	93667
	CEW07F.204	93665		CEW07F.198	93666		CEW07F.221	93667
	CEW07F.209	93665		CEW07F.199	93666		CEW07F.222	93671
	CEW07F.210	93665		CEW07F.200	93666		CEW07F.223	93671
	CEW07F.211	93665		CEW07F.201	93666		CEW07F.224	93671
	CEW07F.212	93665		CEW07F.202	93666		CEW07F.225	93671
	CEW07F.215	93665		CEW07F.203	93666		CEW07F.226	93671
	CEW07F.216	93665		CEW07F.214	93666		CEW07F.227	93671
	CEW07F.218	93665		CEW07F.233	93670		CEW07F.231	93671
	CEW07F.228	93668		CEW07F.234	93670		CEW07F.232	93671
	CEW07F.229	93668		CEW07F.235	93670		CEW07F.244	93671
	CEW07F.230	93668		CEW07F.236	93670		CEW07F.245	93671
	CEW07F.238	93668		CEW07F.237	93670		CEW07F.246	93671
	CEW07F.239	93668		CEW07F.240	93670		CEW07F.247	93671
	CEW07F.241	93668		CEW07F.253	93670		CEW07F.248	93671
	CEW07F.242	93668		CEW07F.254	93670		CEW07F.249	93671
	CEW07F.243	93668		CEW07F.255	93670		CEW07F.250	93671
	CEW07F.102	93659		CEW07F.262	93673		CEW07F.251	93671
	CEW07F.103	93659		CEW07F.268	93673		CEW07F.252	93671
	CEW07F.105	93659		CEW07F.269	93673		CEW07F.256	93671
	CEW07F.107	93659		CEW07F.284	93676		CEW07F.257	93671
	CEW07F.108	93659		CEW07F.287	93676		CEW07F.258	93671
	CEW07F.111	93659		CEW07F.289	93676		CEW07F.259	93671
	CEW07F.112	93659		CEW07F.297	93676		CEW07F.260	93671
	CEW07F.113	93659		CEW07F.298	93676		CEW07F.101	93660
	CEW07F.114	93659		CEW07F.299	93676		CEW07F.104	93660
	CEW07F.116	93659		CEW07F.301	93676		CEW07F.106	93660
	CEW07F.118	93659		CEW07F.307	93676		CEW07F.110	93660
	CEW07F.119	93659		CEW07F.320	93682		CEW07F.129	93660
	CEW07F.122	93659		CEW07F.321	93682		CEW07F.130	93660
	CEW07F.123	93659		CEW07F.322	93682		CEW07F.131	93660
	CEW07F.125	93659		CEW07F.323	93682		CEW07F.132	93660
	CEW07F.127	93659		CEW07F.324	93682		CEW07F.133	93660

CEW07F.134	93659	CEW07F.325	93682	CEW07F.154	93661
CEW07F.135	93659	CEW07F.326	93682	CEW07F.155	93661
CEW07F.136	93659	CEW07F.327	93682	CEW07F.156	93661
CEW07F.137	93659	CEW07F.328	93682	CEW07F.157	93661
CEW07F.138	93659	CEW07F.330	93682	CEW07F.158	93661
CEW07F.139	93659	CEW05F.077	90924	CEW07F.160	93661
CEW07F.140	93659	CEW05F.079	90924	CEW07F.161	93661
CEW07F.141	93659	CEW05F.081	90924	CEW07F.162	93661
CEW07F.142	93659	CEW05F.087	90924	CEW07F.163	93661
CEW07F.143	93659	CEW05F.215	90941	CEW07F.164	93661
CEW07F.144	93659	CEW05F.219	90941	CEW07F.165	93661
CEW07F.145	93659	CEW05F.221	90941	CEW07F.166	93661
CEW07F.146	93659	CEW05F.224	90941	CEW07F.167	93661
CEW07F.149	93659	CEW05F.225	90941	CEW07F.168	93661
CEW07F.150	93659	CEW05F.231	90941	CEW07F.169	93661
CEW07F.261	93672	CEW05F.233	90941	CEW07F.263	93674
CEW07F.264	93672	CEW07F.333	93684	CEW07F.266	93674
CEW07F.265	93672	CEW07F.335	93684	CEW07F.267	93674
CEW07F.277	93675	CEW07F.336	93684	CEW07F.270	93674
CEW07F.278	93675	CEW07F.337	93684	CEW07F.271	93677
CEW07F.280	93675	CEW07F.338	93684	CEW07F.272	93677
CEW07F.281	93675	CEW07F.339	93684	CEW07F.273	93677
CEW07F.282	93675	CEW07F.340	93684	CEW07F.274	93677
CEW07F.283	93675	CEW07F.345	93684	CEW07F.275	93677
CEW07F.288	93675	CEW05F.091	90937	CEW07F.276	93677
CEW07F.290	93675	CEW05F.098	90937	CEW07F.279	93677
CEW07F.291	93675	CEW05F.100	90937	CEW07F.285	93677
CEW07F.292	93675	CEW05F.102	90937	CEW07F.286	93677
CEW07F.293	93675	CEW05F.106	90937	CEW07F.300	93677
CEW07F.294	93675	CEW05F.107	90937	CEW07F.303	93677
CEW07F.295	93675	CEW05F.114	90937	CEW07F.310	93677
CEW07F.302	93675	CEW05F.320	90935	CEW07F.312	93677
CEW07F.304	93675	CEW05F.321	90935	CEW07F.313	93677
CEW07F.305	93675	CEW05F.322	90935	CEW07F.314	93677
CEW07F.308	93675	CEW05F.325	90935	CEW07F.316	93677
CEW07F.309	93675	CEW05F.326	90935	CEW07F.317	93677
CEW07F.311	93675	CEW05F.328	90935	CEW07F.318	93677
CEW07F.315	93675	CEW05F.329	90935	CEW07F.319	93677
CEW05F.075	90921	CEW07F.060	93680	CEW05F.078	90922
CEW05F.076	90921	CEW07F.062	93680	CEW05F.080	90922
CEW05F.084	90921	CEW07F.066	93680	CEW05F.083	90922
CEW05F.086	90921	CEW05F.243	90966	CEW05F.088	90922
CEW05F.089	90921	CEW05F.245	90966	CEW05F.206	90942
CEW05F.090	90921	CEW05F.246	90966	CEW05F.207	90942
CEW05F.208	90940	CEW05F.249	90966	CEW05F.209	90942

CEW05F.213	90940	CEW05F.252	90966	CEW05F.210	90942
CEW05F.217	90940	CEW05F.255	90966	CEW05F.211	90942
CEW05F.223	90940	CEW05F.256	90966	CEW05F.212	90942
CEW05F.226	90940	CEW05F.262	90966	CEW05F.214	90942
CEW05F.227	90940	CEW05F.265	90966	CEW05F.216	90942
CEW05F.228	90940	CEW05F.266	90966	CEW05F.218	90942
CEW05F.229	90940	CEW05F.267	90966	CEW05F.236	90942
CEW05F.230	90940	CEW05F.269	90966	CEW05F.237	90942
CEW05F.232	90940	CEW05F.272	90966	CEW05F.238	90942
CEW05F.234	90940	CEW05F.274	90966	CEW05F.239	90942
CEW05F.235	90940	CEW05F.275	90966	CEW05F.240	90942
CEW07F.331	93683	CEW05F.277	90966	CEW05F.241	90942
CEW07F.332	93683	CEW05F.279	90966	CEW05F.242	90942
CEW07F.344	93683	CEW05F.280	90966	CEW07F.346	93686
CEW07F.352	93683	CEW05F.281	90966	CEW07F.347	93686
CEW05F.057	90939	CEW05F.282	90966	CEW07F.348	93686
CEW05F.059	90939	CEW05F.283	90966	CEW07F.350	93686
CEW05F.060	90939	CEW05F.284	90966	CEW07F.351	93686
CEW05F.061	90939	CEW05F.286	90966	CEW07F.353	93686
CEW05F.062	90939	CEW05F.289	90966	CEW07F.354	93686
CEW05F.063	90939	CEW05F.290	90966	CEW07F.355	93686
CEW05F.064	90939	CEW05F.018	90944	CEW07F.356	93686
CEW05F.066	90939	CEW05F.023	90944	CEW07F.357	93686
CEW05F.069	90939	CEW05F.025	90944	CEW05F.058	90938
CEW05F.070	90939	CEW05F.028	90944	CEW05F.065	90938
CEW05F.072	90939	CEW05F.029	90944	CEW05F.067	90938
CEW05F.073	90939	CEW05F.032	90944	CEW05F.068	90938
CEW05F.092	90939	CEW05F.035	90944	CEW05F.071	90938
CEW05F.093	90939	CEW05F.036	90944	CEW05F.074	90938
CEW05F.094	90939	CEW05F.037	90944	CEW05F.105	90938
CEW05F.095	90939	CEW05F.039	90944	CEW05F.108	90938
CEW05F.096	90939	CEW05F.042	90944	CEW05F.109	90938
CEW05F.097	90939	CEW05F.048	90944	CEW05F.110	90938
CEW05F.101	90939	CEW05F.049	90944	CEW05F.111	90938
CEW05F.103	90939	CEW05F.299	90926	CEW05F.113	90938
CEW05F.104	90939	CEW05F.300	90926	CEW05F.115	90938
CEW05F.112	90939	CEW05F.301	90926	CEW05F.323	90936
CEW05F.116	90939	CEW05F.302	90926	CEW05F.324	90936
CEW05F.117	90939	CEW05F.303	90926	CEW05F.327	90936
CEW05F.118	90939	CEW05F.319	90926	CEW05F.330	90936
CEW07F.055	93695	CEW05F.335	90926	CEW05F.331	90936
CEW07F.056	93695	CEW07F.074	93696	CEW05F.332	90936
CEW07F.058	93695	CEW07F.077	93696	CEW05F.333	90936
CEW07F.063	93695	CEW07F.079	93696	CEW05F.334	90936
CEW07F.072	93695	CEW07F.080	93696	CEW07F.057	93699

CEW05F.001	90949	CEW07F.084	93696	CEW07F.061	93699
CEW05F.003	90949	CEW07F.085	93696	CEW07F.065	93699
CEW05F.008	90949	CEW07F.088	93696	CEW07F.067	93699
CEW05F.011	90949	CEW05F.133	90957	CEW07F.071	93699
CEW05F.122	90949	CEW05F.134	90957	CEW05F.004	90951
CEW05F.125	90949	CEW05F.135	90957	CEW05F.005	90951
CEW05F.129	90949	CEW05F.141	90957	CEW05F.006	90951
CEW05F.130	90949	CEW05F.143	90957	CEW05F.007	90951
CEW05F.247	90948	CEW05F.144	90957	CEW05F.009	90951
CEW05F.248	90948	CEW05F.149	90957	CEW05F.010	90951
CEW05F.250	90948	CEW05F.151	90957	CEW05F.012	90951
CEW05F.251	90948	CEW05F.153	90957	CEW05F.013	90951
CEW05F.253	90948	CEW05F.156	90957	CEW05F.119	90951
CEW05F.257	90948	CEW05F.157	90957	CEW05F.120	90951
CEW05F.258	90948	CEW05F.158	90957	CEW05F.121	90951
CEW05F.260	90948	CEW05F.159	90957	CEW05F.123	90951
CEW05F.263	90948	CEW05F.160	90957	CEW05F.124	90951
CEW05F.264	90948	CEW05F.163	90957	CEW05F.126	90951
CEW05F.268	90948	CEW05F.172	90957	CEW05F.128	90951
CEW05F.271	90948	CEW05F.181	90957	CEW05F.131	90951
CEW05F.273	90948	CEW05F.182	90957	CEW05F.259	90953
CEW05F.276	90948	CEW05F.185	90957	CEW05F.287	90953
CEW05F.278	90948	CEW05F.187	90957	CEW05F.288	90953
CEW05F.285	90948	CEW05F.191	90957	CEW05F.293	90953
CEW05F.292	90948	CEW05F.192	90957	CEW07F.098	93706
CEW07F.100	93704			CEW07F.501	93706
CEW07F.502	93704			CEW07F.509	93706
CEW07F.507	93704			CEW07F.516	93706
CEW07F.510	93704			CEW07F.518	93706
CEW07F.513	93704			CEW07F.519	93706
CEW07F.514	93704			CEW05F.014	90945
CEW07F.515	93704			CEW05F.017	90945
CEW05F.015	90943			CEW05F.019	90945
CEW05F.016	90943			CEW05F.021	90945
CEW05F.020	90943			CEW05F.022	90945
CEW05F.026	90943			CEW05F.040	90945
CEW05F.027	90943			CEW05F.041	90945
CEW05F.030	90943			CEW05F.046	90945
CEW05F.031	90943			CEW05F.305	90952
CEW05F.033	90943			CEW05F.306	90952
CEW05F.034	90943			CEW05F.308	90952
CEW05F.038	90943			CEW05F.309	90952
CEW05F.043	90943			CEW05F.310	90952
CEW05F.044	90943			CEW05F.311	90952
CEW05F.045	90943			CEW05F.312	90952

CEW05F.047	90943	CEW05F.313	90952
CEW05F.051	90943	CEW05F.314	90952
CEW05F.052	90943	CEW05F.315	90952
CEW05F.053	90943	CEW05F.316	90952
CEW05F.054	90943	CEW05F.317	90952
CEW05F.055	90943	CEW05F.318	90952
CEW05F.056	90943	CEW07F.078	93702
CEW05F.304	90925	CEW07F.087	93702
CEW05F.307	90925	CEW07F.090	93702
CEW07F.075	93701	CEW07F.093	93702
CEW07F.076	93701	CEW07F.094	93702
CEW07F.082	93701	CEW07F.095	93702
CEW07F.083	93701	CEW05F.132	90950
CEW07F.091	93701	CEW05F.138	90950
CEW05F.136	90958	CEW05F.139	90950
CEW05F.137	90958	CEW05F.142	90950
CEW05F.140	90958	CEW05F.145	90950
CEW05F.148	90958	CEW05F.147	90950
CEW05F.150	90958	CEW05F.155	90950
CEW05F.152	90958	CEW05F.162	90950
CEW05F.154	90958	CEW05F.173	90950
CEW05F.161	90958	CEW05F.174	90950
CEW05F.164	90958	CEW05F.177	90950
CEW05F.165	90958	CEW05F.194	90950
CEW05F.166	90958	CEW05F.195	90950
CEW05F.167	90958	CEW05F.196	90950
CEW05F.168	90958	CEW05F.197	90950
CEW05F.169	90958	CEW05F.198	90950
CEW05F.170	90958	CEW05F.199	90950
CEW05F.171	90958	CEW05F.200	90950
CEW05F.175	90958	CEW05F.201	90950
CEW05F.176	90958	CEW05F.202	90950
CEW05F.178	90958	CEW05F.203	90950
CEW05F.179	90958	CEW05F.204	90950
CEW05F.180	90958	CEW05F.205	90950
CEW05F.183	90958		
CEW05F.184	90958		
CEW05F.186	90958		
CEW05F.188	90958		
CEW05F.189	90958		
CEW05F.190	90958		
CEW05F.193	90958		
CEW05F.353	90958		

APPENDIX 2

Supplementary Information for Chapter 2

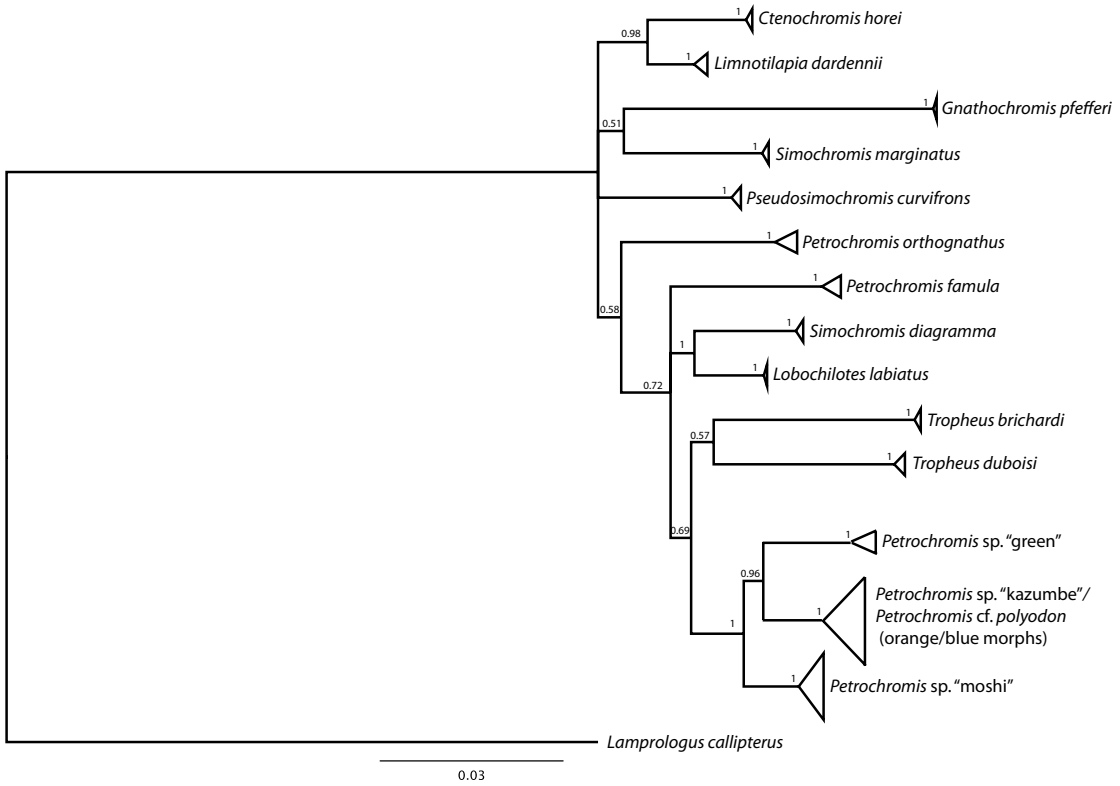


Figure S1. Phylogenetic relationships among tropheine cichlids from the Kigoma region based on cytB and ND2 sequences. Numbers on the nodes are posterior probabilities from the Bayesian phylogenetic analysis. All species are strongly supported as reciprocally monophyletic except for orange and blue morph *Petrochromis*. Four of the six *Petrochromis* species present in the Kigoma region form a strongly supported clade: *P.* sp. "moshi", *P.* sp. "green" and the orange and blue *Petrochromis* morphs which are the focus of this study (see also Figure 2.1a).

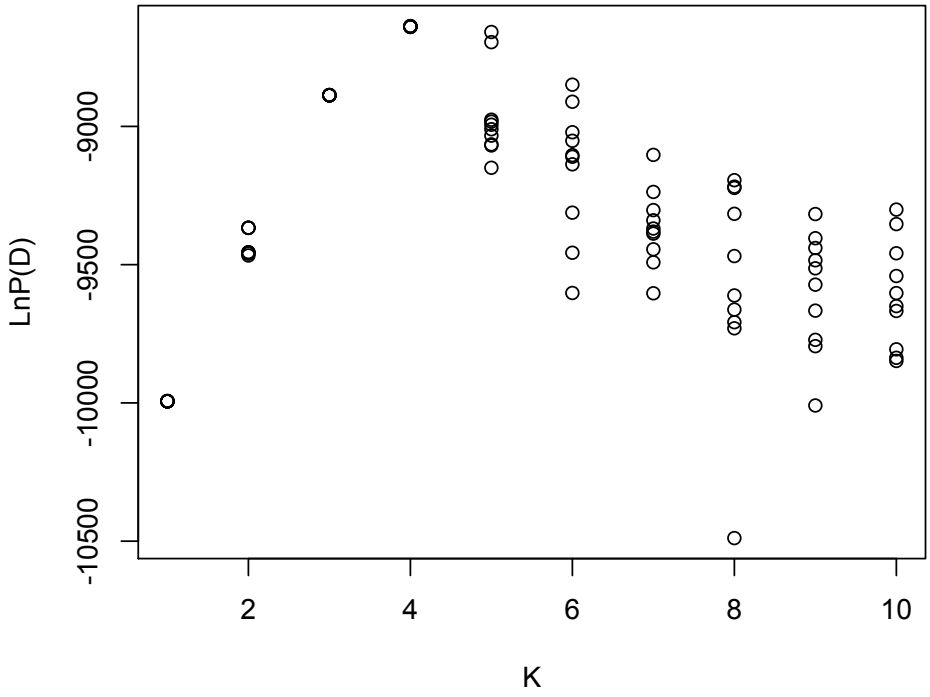


Figure S2. LnP(D) values from STRUCTURE over 10 runs each at 10 values of *K*. LnP(D) is maximized at $K = 4$, and thereafter variance among runs increases substantially.

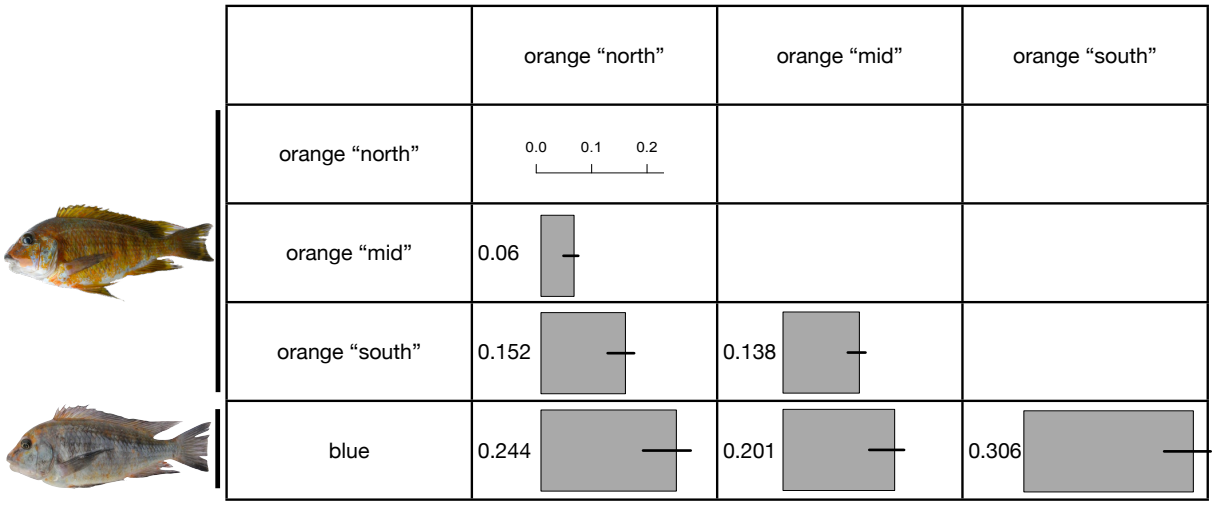


Figure S3. Pairwise F_{ST} values from populations identified in STRUCTURE runs.

Table S1. Species studied and specimen information. Specimens are retained as vouchers at the Cornell Museum of Vertebrates (CUMV); accession numbers are noted in this column. Cells with "x" indicate that data for this specimen and genetic locus was collected, but is not available in a genetic repository at the time of publication of this document.

Species	Sp. Authority	Individual ID	CUMV	GenBank ND2	GenBank CytB	Genotyped in this study
Ctenochromis horei	Günther 1894	CEW07F.051	93698	x	x	
		CEW07F.052	93698	x	x	
Gnathochromis pfefferi	Boulenger 1898	PBMLT02.068	88652	EF679245	EF679277	
	Boulenger 1898	PBMLT02.224	88652	x	x	
Lamprologus callipterus	Boulenger 1906	PBMLT02.083	88650	EF679240	EF679272	
Limnotilapia dardennii	Boulenger 1899	PBMLT02.264	88642	x	x	
	Boulenger 1899	PBMLT02.265	88642	EF679249	EF679281	
Lobochilotes labiatus	Boulenger 1898	PBMLT02.302	88643	x	x	
	Boulenger 1898	PBMLT02.303	88643	EF679250	EF679282	
Petrochromis cf. *polyodon*[1]	Boulenger 1898	CEW05F.002	90949	x		x
		CEW05F.050	90943	x		x
		CEW05F.099	90939			x
		CEW05F.127	90949	x		x
		CEW05F.244	90948	x		x
		CEW05F.261	90948	x		x
		CEW05F.295	90930	x	x	x
		CEW07F.109	93659	x		x
		CEW07F.115	93659	x		x
		CEW07F.117	93659	x		x
		CEW07F.121	93659			x
		CEW07F.124	93659			x
		CEW07F.126	93659			x
		CEW07F.128	93659			x
		CEW07F.170	93663	x		x
		CEW07F.172	93663	x		x
		CEW07F.176	93663	x		x
		CEW07F.178	93663	x		x
		CEW07F.189	93663			x
		CEW07F.191	93663			x
		CEW07F.296	93675			x
		CEW07F.306	93675	x		x
		CEW07F.329	93678	x		x
		CEW07F.342	93683	x		x
		PBMLT02.253	88725	x	x	
Petrochromis famula	Matthes & Trewavas 1960	PBMLT02.212	88724	x	x	
		PBMLT02.213	88724	EF679265	EF679297	
Petrochromis sp. "moshi"[2]	undescribed	CEW05F.182	90957	x	x	
		CEW05F.185	90957	x	x	
		PBMLT02.011	88710	x	x	
		PBMLT02.073	88710	EF679256	EF679288	
Petrochromis orthognathus	Matthes 1959	PBMLT02.137	88722	x	x	
		PBMLT02.138	88722	x	x	
Petrochromis sp. "green"[3]	undescribed	CEW05F.024	NA			x
		CEW05F.082	90921			x
		CEW05F.296	90929	x	x	x
		CEW05F.297	90929	x	x	x
Petrochromis sp. "kazumbe"[4]	undescribed	CEW05F.001	90949	x		x
		CEW05F.003	90949	x		x
		CEW05F.008	90949			x
		CEW05F.011	90949			x
		CEW05F.015	90943	x		x
		CEW05F.016	90943	x		x
		CEW05F.020	90943			x
		CEW05F.026	90943			x
		CEW05F.027	90943			x

Species	Sp. Authority	Individual ID	CUMV	GenBank ND2	GenBank CytB	Genotyped in this study
		CEW05F.030	90943			x
		CEW05F.031	90943			x
		CEW05F.033	90943			x
		CEW05F.034	90943			x
		CEW05F.038	90943			x
		CEW05F.043	90943			x
		CEW05F.044	90943			x
		CEW05F.045	90943			x
		CEW05F.047	90943			x
		CEW05F.051	90943			x
		CEW05F.052	90943			x
		CEW05F.053	90943			x
		CEW05F.054	90943			x
		CEW05F.055	90943			x
		CEW05F.056	90943			x
		CEW05F.057	90939	x		x
		CEW05F.059	90939	x		x
		CEW05F.060	90939			x
		CEW05F.061	90939			x
		CEW05F.062	90939			x
		CEW05F.063	90939			x
		CEW05F.064	90939			x
		CEW05F.066	90939			x
		CEW05F.069	90939			x
		CEW05F.070	90939			x
		CEW05F.072	90939			x
		CEW05F.073	90939			x
		CEW05F.075	90921	x		x
		CEW05F.076	90921	x		x
		CEW05F.084	90921			x
		CEW05F.086	90921			x
		CEW05F.089	90921			x
		CEW05F.090	90921			x
		CEW05F.092	90939			x
		CEW05F.093	90939			x
		CEW05F.094	90939			x
		CEW05F.095	90939			x
		CEW05F.096	90939			x
		CEW05F.097	90939			x
		CEW05F.101	90939			x
		CEW05F.103	90939			x
		CEW05F.104	90939			x
		CEW05F.112	90939			x
		CEW05F.116	90939			x
		CEW05F.117	90939			x
		CEW05F.118	90939			x
		CEW05F.122	90949			x
		CEW05F.125	90949			x
		CEW05F.129	90949			x
		CEW05F.130	90949			x
		CEW05F.136	90958	x		x
		CEW05F.137	90958	x		x
		CEW05F.140	90958			x
		CEW05F.148	90958			x
		CEW05F.150	90958			x
		CEW05F.152	90958			x
		CEW05F.154	90958			x
		CEW05F.161	90958			x
		CEW05F.164	90958			x
		CEW05F.165	90958			x

Species	Sp. Authority	Individual ID	CUMV	GenBank ND2	GenBank CytB	Genotyped in this study
		CEW05F.166	90958			x
		CEW05F.167	90958			x
		CEW05F.168	90958			x
		CEW05F.169	90958			x
		CEW05F.170	90958			x
		CEW05F.171	90958			x
		CEW05F.175	90958			x
		CEW05F.176	90958			x
		CEW05F.178	90958			x
		CEW05F.179	90958			x
		CEW05F.180	90958			x
		CEW05F.183	90958	x		x
		CEW05F.184	90958	x	x	x
		CEW05F.186	90958	x	x	x
		CEW05F.188	90958			x
		CEW05F.189	90958			x
		CEW05F.190	90958			x
		CEW05F.193	90958			x
		CEW05F.208	90940			x
		CEW05F.213	90940			x
		CEW05F.217	90940			x
		CEW05F.223	90940			x
		CEW05F.226	90940			x
		CEW05F.227	90940			x
		CEW05F.228	90940			x
		CEW05F.229	90940			x
		CEW05F.230	90940			x
		CEW05F.232	90940			x
		CEW05F.234	90940			x
		CEW05F.235	90940			x
		CEW05F.247	90948			x
		CEW05F.248	90948			x
		CEW05F.250	90948			x
		CEW05F.251	90948			x
		CEW05F.253	90948			x
		CEW05F.257	90948			x
		CEW05F.258	90948			x
		CEW05F.260	90948			x
		CEW05F.263	90948			x
		CEW05F.264	90948			x
		CEW05F.268	90948			x
		CEW05F.271	90948			x
		CEW05F.273	90948			x
		CEW05F.276	90948			x
		CEW05F.278	90948			x
		CEW05F.285	90948			x
		CEW05F.292	90948			x
		CEW05F.304	90925			x
		CEW05F.307	90925			x
		CEW07F.055	93695			x
		CEW07F.056	93695			x
		CEW07F.058	93695			x
		CEW07F.063	93695			x
		CEW07F.072	93695			x
		CEW07F.075	93701			x
		CEW07F.076	93701			x
		CEW07F.082	93701			x
		CEW07F.083	93701			x
		CEW07F.091	93701			x
		CEW07F.100	93704			x

Species	Sp. Authority	Individual ID	CUMV	GenBank ND2	GenBank CytB	Genotyped in this study
		CEW07F.102	93659	x		x
		CEW07F.103	93659	x		x
		CEW07F.105	93659			x
		CEW07F.107	93659			x
		CEW07F.108	93659			x
		CEW07F.111	93659			x
		CEW07F.112	93659			x
		CEW07F.113	93659			x
		CEW07F.114	93659			x
		CEW07F.116	93659			x
		CEW07F.118	93659			x
		CEW07F.119	93659			x
		CEW07F.122	93659			x
		CEW07F.123	93659			x
		CEW07F.125	93659			x
		CEW07F.127	93659			x
		CEW07F.134	93659			x
		CEW07F.135	93659			x
		CEW07F.136	93659			x
		CEW07F.137	93659			x
		CEW07F.138	93659			x
		CEW07F.139	93659			x
		CEW07F.140	93659			x
		CEW07F.141	93659			x
		CEW07F.142	93659			x
		CEW07F.143	93659			x
		CEW07F.144	93659			x
		CEW07F.145	93659			x
		CEW07F.146	93659			x
		CEW07F.149	93659			x
		CEW07F.150	93659			x
		CEW07F.171	93663	x		x
		CEW07F.173	93663	x		x
		CEW07F.175	93663			x
		CEW07F.180	93663			x
		CEW07F.183	93663			x
		CEW07F.185	93663			x
		CEW07F.187	93663			x
		CEW07F.190	93663			x
		CEW07F.204	93665			x
		CEW07F.209	93665			x
		CEW07F.210	93665			x
		CEW07F.211	93665			x
		CEW07F.212	93665			x
		CEW07F.215	93665			x
		CEW07F.216	93665			x
		CEW07F.218	93665			x
		CEW07F.228	93668			x
		CEW07F.229	93668			x
		CEW07F.230	93668			x
		CEW07F.238	93668			x
		CEW07F.239	93668			x
		CEW07F.241	93668			x
		CEW07F.242	93668			x
		CEW07F.243	93668			x
		CEW07F.261	93672	x		x
		CEW07F.264	93672	x		x
		CEW07F.265	93672			x
		CEW07F.277	93675			x
		CEW07F.278	93675			x

Species	Sp. Authority	Individual ID	CUMV	GenBank ND2	GenBank CytB	Genotyped in this study
		CEW07F.280	93675			x
		CEW07F.281	93675			x
		CEW07F.282	93675			x
		CEW07F.283	93675			x
		CEW07F.288	93675			x
		CEW07F.290	93675			x
		CEW07F.291	93675			x
		CEW07F.292	93675			x
		CEW07F.293	93675			x
		CEW07F.294	93675			x
		CEW07F.295	93675			x
		CEW07F.302	93675			x
		CEW07F.304	93675			x
		CEW07F.305	93675			x
		CEW07F.308	93675			x
		CEW07F.309	93675			x
		CEW07F.311	93675			x
		CEW07F.315	93675			x
		CEW07F.331	93683			x
		CEW07F.332	93683			x
		CEW07F.344	93683			x
		CEW07F.352	93683			x
		CEW07F.502	93704			x
		CEW07F.507	93704			x
		CEW07F.510	93704			x
		CEW07F.513	93704			x
		CEW07F.514	93704			x
		CEW07F.515	93704			x
		PBMLT02.171	88723	x	x	
		PBMLT02.253	88725	EF679257	EF679289	
Pseudosimochromis curvifrons	Poll 1942	PBMLT02.309	89316	x	x	
		PBMLT02.310	89316	x	x	
Simochromis diagramma	Günther 1894	PBMLT02.256	88719	EF679259	EF679291	
		PBMLT02.257	88719	x	x	
Simochromis marginatus	Poll 1956	PBMLT02.078	88720	x	x	
		PBMLT02.292	88720	x	x	
Tropheus brichardi	Nelissen & Thys van den Audenaerde 1975	PBMLT02.007	88711	EF679262	EF679294	
		PBMLT04.030	90789	x	x	
Tropheus duboisi	Marlier 1959	PBMLT02.124	88651	EF679263	EF679295	
		PBMLT02.345	88651	x	x	

1 The color morph identified as *Petrochromis* cf. *polyodon* in this study ("blue morph") is known by this name in the aquarium trade in the Kigoma region. How it corresponds to the species identified by Boulenger has not been examined. Based sequences available in GenBank, this species needs taxonomic reassessment, particularly with regard to the identity of Boulenger (1898)'s original type description (C. Wagner, unpublished work). It is rare in the Kigoma region compared to the abundance of *P.* sp. "kazumbe".

2 *Petrochromis* sp. "moshi" is well known in the aquarium trade, but undescribed. Based on GenBank sequences and the identifications of specimens associated with these sequences, it appears to have close affinities to the described species *P. ephippium* (C. Wagner, unpublished data), but does not co-occur with this species in the Kigoma region. For additional CUMV voucher information for *P.* moshi genotypes used in this study, see Wagner and McCune 2009.

3 The species that we call *P.* sp. "green" in this study is, to our knowledge, not previously known from the aquarium trade. Its mouth morphology is reminiscent of the described species *P. macrognathus* (Yamaoka, 1983), but based on GenBank sequences identified as this species, and other specimens of *P. macrognathus* that we obtained and sequenced, this species is not closely related to the described *P. macrognathus*. We therefore use the phenotypically descriptive *P.* sp. "green" here to make this distinction.

4 *Petrochromis* sp. "kazumbe" is well known in the aquarium trade, but undescribed. It is abundant in the Kigoma region.

References

Boulenger, G. A. 1898, Report on the fishes recently obtained by Mr. J. E. S. Moore in Lake Tanganyika. Proceedings of the General Meetings for Scientific Business of the Zoological Society of London 1898 pt 3:494-497.

—. 1899. Second contribution to the ichthyology of Lake Tanganyika.-- On the fishes obtained by the Congo Free State Expedition under Lieut. Lemaire in 1898. Transactions of the Zoological Society of London 15:87-96.

—. 1906. Fourth contribution to the ichthyology of Lake Tanganyika.--Report on the collection of fishes made by Dr. W. A. Cunnington during the Third Tanganyika Expedition, 1904-1905. . Transactions of the Zoological Society of London 17:537-619.

Gunther, A. 1894, Descriptions of the reptiles and fishes collected by Mr. E. Coode-Hore on Lake Tanganyika Proceedings of the General Meetings for Scientific Business of the Zoological Society of London 1893 4:628-632.

Marlier, G. 1959. Observations sur la biologie littorale du Lac Tanganyika. Revue de Zoologie et de Botanique Africaines 59:164-183.

Matthes, H. 1959. Un cichlide nouveau du Lac Tanganyika: *Petrochromis orthognathus* n. sp. Revue de Zoologie et Botanique Africaines 60:335-341.

Matthes, H., and E. Trewavas. 1960. *Petrochromis famula* n. sp., a cichlid fish of Lake Tanganyika. Revue de Zoologie et de Botanique Africaines 61:349-357.

Nelissen, M. H. J., and D. F. E. T. v. d. Audenaerde. 1975. Description of *Tropheus brichardi* sp. nov. from Lake Tanganyika (Pisces, Cichlidae). Revue de Zoologique Africaine 89:974-980.

Poll, M. 1942. Cichlidae nouveaux du lac Tanganika appartenant aux collections du Musée du Congo. Revue de Zoologie et Botanique Africaines 36:343-360.

—. 1956. Poissons Cichlidae. Exploration Hydrobiologique du Lac Tanganika (1946-1947), resultats scientifique 3:1-619.

Yamaoka, K. 1983. A Revision Of The Cichlid Fish Genus Petrochromis From Lake Tanganyika, With Description Of A New Species. Japanese Journal Of Ichthyology 30:129-141.

APPENDIX 3

Supplementary Information for Chapter 3

1. Phylogenetic analyses

1.1 Taxon sampling and sequencing

Species sampled and GenBank numbers associated with gene sequences used in phylogenetic analyses are listed in Supplementary Table 1.

1.2 Phylogenetic analyses

We assembled the genetic dataset using functions in the R packages APE (Paradis et al. 2004) and Phyloch (Heibl 2010), and aligned the dataset with MAFFT (Katoh et al. 2009).

We used RAxML for phylogenetic analyses (Stamatakis 2006). We partitioned the dataset by gene, using a GTR+gamma model of sequence evolution for each gene partition. We completed a full maximum likelihood search and 100 bootstrap replicates of RAxML's rapid bootstrap algorithm (Stamatakis et al. 2008). To account for uncertainty in branch length estimates as well as topology, we estimated branch lengths for each bootstrap replicate topology in RAxML, giving a total of 101 trees with topology and branch length estimates.

To ultrametricize and time-calibrate this set of trees, we used PATHd8 (Britton et al. 2007). We used four geological dates to time-calibrate the trees. Two of these dates were associated with the breakup of Gondwana: the African-Madagascar split (121-165 million years ago), placed at the node representing the common ancestor of mainland and Madagascar cichlids; the Madagascar-India split (63-88 million years ago), placed at the node representing the common ancestor of Indian and Madagascar plus African mainland cichlids (Genner et al. 2007). We also included the age of the earliest known fossil *Oreochromis* (6 million years; Carnevale et al. 2003), placed at the node representing the common ancestor of *Oreochromis* and *Sarotherodon* (*Oreochromis* and *Sarotherodon* cannot be distinguished based on fossilized characters, thus this placement is conservative). Additionally, we used the age of Lake Nabugabo, 5000 years (Genner et al. 2007), as a recent calibration point. Because cichlids from Lake Nabugabo are not reciprocally monophyletic, we applied this divergence time to the node representing each Nabugabo species and its most recent common ancestor, repeating this procedure for each of the four Nabugabo species included in the tree, and replicating this procedure over the set of 101 trees. We then drew 95% confidence intervals on node ages from the distribution of branching times estimated from these sets of calibrated ultrametric trees.

The best maximum likelihood topology from a full RAxML search, with bootstrap values from 100 rounds of bootstrapping, is provided in Supplementary Figure 1.

1.3 Tree manipulation for phylogenetic logistic regression

To incorporate our phylogeny into phylogenetic logistic regression analyses, we trimmed the single best ML tree to include a) only lineages that occur in lakes, b) a single taxon for each lake in which cichlids have diversified. For lineages present in multiple lakes, we added a tip to the tree for each instance where the lineage is found in a unique lake, such that each lineage found in multiple lakes is represented as a polytomy with a tip corresponding to each lake where it is present. We set branch lengths on these added tips to have a total length that matched that expected under a pure birth model.

2. Diversification state, trait and environmental data

2.1 Data included and source information

Lake physical and environmental variables, are reported in Supplementary Table 2. Lake surface areas that were not reported in the literature were measured using distance-calibrated Google Earth satellite images and the software ImageJ (Abramoff et al. 2004). Net solar radiation is the difference between the influx of solar radiation and the reflectance of heat energy back into space. We used the average of monthly values from 2010, obtained from the NASA Langley Research Center Atmospheric Science Data Center.

We obtained information about the distribution of large predatory fish of the genera *Lates, Hydrocyanus* and *Hepsetus* from FishBase (Froese and Pauly 2010).

Trait data and diversification states are reported in Supplemenatary Table 3.

2.2 Time for diversification

We calculated “time for diversification” for lineages using either the midpoint of geological age estimates for the lake (either most recent desiccation or basin age, if no evidence for desiccation exists) or the mean stem age of the radiating group estimated from our calibrated molecular phylogenies. If both ages were available we used the geological age of the lake, with the exception of lineages in Lake Tanganyika, where radiating groups differ substantially in age. Previous work has used similar approaches, combining geological and molecular genetic information to assess relative tempos of speciation (McCune 1997; McCune and Lovejoy 1998). Supplementary Table 4 provides the times for diversification that we inferred for all lineages in the dataset, and references information for geologically-based dates.

Because of potential error arising from combining molecular phylogenetic estimates of clade age with geologic dates, we also did analyses using only geologically-based lake ages (see SI section 4 below). The results of these analyses are qualitatively identical to those using time for diversification.

2.3 Diversification "thresholds"

We coded each lineage in each lake as one of two diversification states – "diversifying" or "nondiversifying" using 2 different thresholds to identify diversifying lineages. At the lowest threshold ("threshold 1"), we identified any lineage that had undergone at least one intralacustrine speciation event. Under this criterion, any lineage that had at least one endemic species in a lake co-occurring with its sister taxon (either a widespread species or a lake endemic itself) would be coded as diversifying. Single endemic species not co-occurring with a sister taxon were not considered to be diversifying. All results in the main text are for threshold 1. As an additional test, we coded lineages as diversifying only if they had produced at least 5 endemic species within a given lake ("threshold 5"). For results of analyses conduced for at this higher diversification threshold, see sections 4 and 5 below.

2.4. Treating radiation as a binary variable

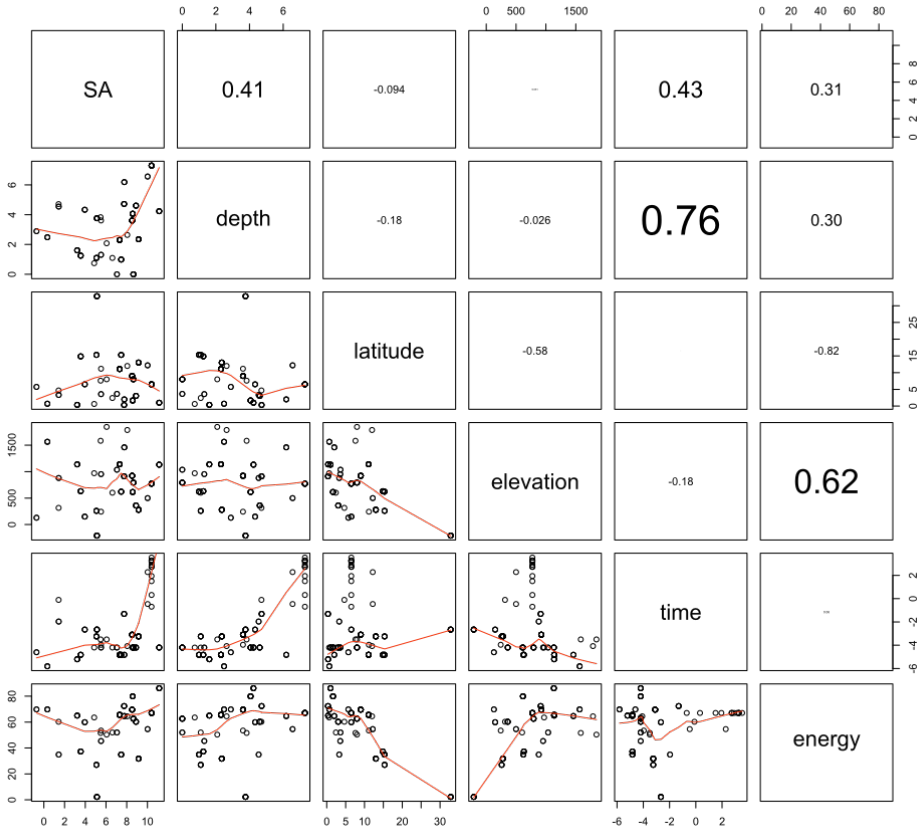
In this study, we focus on explaining the presence and absence of diversification; not the species richness of diversifying lineages. There are two major reasons we made the decision to treat the data this way: 1. We here ask if there are lineage-specific or environmental properties that promote intralacustrine diversification. This simply a different question than asking what determines the species richness of diversifying lineages; these are each valid and interesting questions. 2. Because many lineages have colonized lakes and have not subsequently diversified, the species richness dataset including all colonizing lineages would be highly skewed towards 0-values; the response variable therefore has a strongly zero-skewed distribution. Therefore, the binary framework we use here is more analytically appropriate for the dataset.

3. Correlation between predictor variables

We checked for collinearity between predictor variables prior to including variables together in multiple regression models. We calculated Pearson correlation coefficients (r^2) for all pairs of continuous predictor variables (see Supplementary Figure 2). Among continuous predictor variables, lake depth and time for diversification were strongly positively correlated ($r^2 = 0.76$), and latitude and environmental energy were strongly negatively correlated ($r^2 = -0.82$).

The collinearity between lake depth and time for diversification is not unexpected, as deeper lakes are generally older because they are less sensitive to climate-driven desiccation. Greater depth could influence diversification by increasing lake stability, and/or through increased habitat dimensionality. To examine the relative explanatory power of depth versus time, we excluded lakes deeper than 150 meters from the dataset, thereby reducing collinearity between depth and time ($r^2 = 0.25$). We then compared models incorporating time, depth, and depth + time as predictors of diversification state. We find that depth alone predicts diversification better than does time alone (ΔAIC 2.996) or depth + time (ΔAIC 1.696). We therefore included depth alone in multiple regression models presented in the main text. Other approaches gave qualitatively similar results.

Because of high collinearity between environmental energy and latitude (r^2 = -0.82), we included the residuals of the linear regression of latitude as a function of net radiation, instead of raw latitude, in multiple regression models. This approach allowed us to ask whether variation in latitude other than that explained by differences in available environmental energy influences cichlid diversification. The residuals of latitude were not strong predictors of diversification in any model set in multiple regression analyses, but excluding them as a predictor variable produced qualitatively identical results.



Supplementary Figure 2. Correlation between continuous predictor variables. There is strong correlation between lake depth and time for diversification (r^2 = 0.76) and between environmental energy and latitude (r^2 = -0.82).

For binary predictor variables, we used the r^2 equivalent suggested by Menard (2000), r^2_L, as an assessment of collinearity. This metric is based on the likelihood of the model with only the intercept (L_O) relative to the model with the predictor variables included (L_M), where

$r^2_L = 1 - \ln(L_O)/\ln(L_M)$.

We removed one variable from each pair of predictor variables with r^2 (or r^2_L) of greater than 0.6 after testing models including variables with correlations higher than this value proved to cause analytical problems (inflations of standard error in parameter estimation, a diagnostic of collinearity problems in logistic regression (Quinn and Keough 2002)). Supplementary Table 5 provides r^2_L for all pairs of binary predictor variables. Mouthbrooding and polygamous mating systems were the only pair of variables with r^2_L greater than 0.6, so we removed mouthbrooding from the multiple regression models shown in the main text.

Supplementary Table 5. Correlation between binary predictor variables, measured as r^2_L, the likelihood-based equivalent of r^2 for binary variables. Mouthbrooding and polygamous mating systems are significantly correlated.

	predators	polygamy	eggspots	haplo eggspots	sexual dichromatism	mouthbrooding
predators		0.106	0.016	0.022	0.054	0.063
polygamy	-		0.271	0.185	0.409	**0.823**
eggspots	-	+		0.560	0.261	0.215
haplo eggspots	-	+	+		0.165	0.164
sexual dichromatism	-	+	+	+		0.297

4. Single predictor variable analyses

We evaluated the relationships between single predictor variables and diversification state using phylogenetic logistic regression (see Methods). The strongest associations with diversification among extrinsic factors were for lake depth, lake age, and time for diversification (see Supplementary Table 6). Sexual dichromatism is also significantly associated with diversification. Because Lake Tanganyika is an outlier in terms of depth and age (it is more than twice as old and deep as any other lake in the dataset), we ran models both with and without lineages present in that lake. Without Tanganyika, depth and time remain significant predictors or radiation, although the strength of these associations decreases. Additionally, surface area becomes a significant negative predictor of diversification. There are also marginally significant negative effects of latitude and the presence of predators on diversification. Sexual dichromatism remains a significant predictor of diversification.

We repeated these analyses for the alternative diversification “threshold” value outlined in section 2, where we scored lineages as diversifying only if they produced at least 5 species, respectively, within a given lake (Supplementary Table 7). At this diversification threshold, lake depth, lake age and time for diversification were again the variables most strongly associated

with diversification. Additionally, there were significant associations between energy, latitute and mouthbrooding for the complete dataset. When Lake Tanganyika was excluded, lake depth, time for diversification and lake age remained strongly significant predictors of diversification, and egg spots, haplochromine egg spots, and sexual dichromatism were also significantly associated with diversification.

5. Multiple regression models with alternative diversification thresholds

We conducted all multiple regression analyses using alternative diversification "threshold" value outlined in section 2, where we scored lineages as diversifying only if they produced at least 5 species, respectively, within a given lake. Results of these analyses are presented in Supplementary Table 8.

The results for this higher diversification threshold are concordant with those for the lower diversification threshold presented in the main text. Relative importance values decrease for all variables as the threshold increases. This is expected, as the number of instances of observed diversification decreases with increased threshold, and thus our power to observe significant associations decreases. The top predictor variables are consistent across thresholds, with one exception. At threshold 5, lake area is no longer a significant predictor of diversification (for thresholds 1 it is a strong negative predictor of diversification). This indicates that the strength of the negative association between lake area and diversification is driven by radiations of very low species richness (< 5) in very small lakes. One possibility for the negative association between lake area and diversification that we discuss in the main text is sampling bias. Data on species present in very small lakes is rare compared to that for large lakes, and small lakes included in the dataset are frequently those known for their endemic cichlids (e.g. Cameroonian crater lakes; Guineas Sink Hole). Furthermore, although our results demonstrate that the occurrence of intralacustrine speciation is not limited by lake area (see main text), species richness in cichlid radiations is limited by lake area (see Seehausen 2006, Wagner et al. in prep). Therefore, if the small lakes included in the dataset disproportionately represent cases of diversification (more so than larger lakes), and small lakes are constrained in species richness (by virtue of their area), the negative area-diversification relationship would be expected to disappear as species richness threshold increases. This is what we observe.

6. References

Abramoff, M. D., P. J. Magelhaes, and S. J. Ram. 2004. Image processing with ImageJ. Biophotonics International 11:36-42.

Barker, P., and F. Gasse. 2003. New evidence for a reduced water balance in East Africa during the Last Glacial Maximum: implication for model-data comparison. Quaternary Science Reviews 22:823-837.

Beadle, L. C. 1974, The inland waters of tropical Africa. An introduction to tropical limnology. . London, Longman.

Bergner, A. G. N., M. R. Strecker, M. H. Trauth, A. Deino, F. Gasse, P. Blisniuk, and M. Duhnforth. 2009. Tectonic and climatic control on evolution of rift lakes in the Central Kenya Rift, East Africa. Quaternary Science Reviews 28:2804-2816.

Bessems, I., D. Verschuren, J. M. Russell, J. Hus, F. Mees, and B. F. Cumming. 2008. Palaeolinmological evidence for widespread late 18th century drought across equatorial East Africa. Palaeogeography Palaeoclimatology Palaeoecology 259:107-120.

Beuning, K. R. M., M. R. Talbot, and K. Kelts. 1997. A revised 30,000-year paleoclimatic and paleohydrologic history of Lake Albert, East Africa. Palaeogeography Palaeoclimatology Palaeoecology 136:259-279.

Britton, T., C. L. Anderson, D. Jacquet, S. Lundqvist, and K. Bremer. 2007. Estimating divergence times in large phylogenetic trees. Systematic Biology 56:741-752.

Carnevale, G., C. Sorbini, and W. T. Landini. 2003. Oreochromis lorenzoi, a new species of tilapiine cichlid from the late Miocene of central Italy. Journal of Vertebrate Paleontology 23:508-516.

Cohen, A. S., J. R. Stone, K. R. M. Beuning, L. E. Park, P. N. Reinthal, D. Dettman, C. Scholz et al. 2007. Ecological consequences of early Lake Pleistocene megadroughts in tropical Africa. Proceedings Of The National Academy Of Sciences Of The United States Of America 104:16422-16427.

Froese, R., and D. Pauly. 2010. FishBase.

Garcin, Y., A. Junginger, D. Melnick, D. O. Olago, M. R. Strecker, and M. H. Trauth. 2009. Late Pleistocene-Holocene rise and collapse of Lake Suguta, northern Kenya Rift. Quaternary Science Reviews 28:911-925.

Genner, M. J., O. Seehausen, D. H. Lunt, D. A. Joyce, P. W. Shaw, G. R. Carvalho, and G. F. Turner. 2007. Age of cichlids: New dates for ancient lake fish radiations. Molecular Biology And Evolution 24:1269-1282.

Hazan, N., M. Stein, A. Agnon, S. Marco, D. Nadel, J. F. W. Negendank, M. J. Schwab et al. 2005. The late quaternary limnological history of Lake Kinneret (Sea of Galilee), Israel. Quaternary Research 63:60-77.

Heibl, C. 2010.phyloch: Tools for phylogenetics and taxonomy.

Johnson, T. C., K. Kelts, and E. Odada. 2000. The holocene history of Lake Victoria. Ambio 29:2-11.

Katoh, K., G. Asimenos, and H. Toh. 2009. Multiple Alignment of DNA Sequences with MAFFT. Bioinformatics for DNA Sequence Analysis:39-64.

Koblmuller, S., U. K. Schliewen, N. Duftner, K. M. Sefc, C. Katongo, and C. Sturmbauer. 2008. Age and spread of the haplochromine cichlid fishes in Africa. Molecular Phylogenetics and Evolution 49:153-169.

Koch, M., S. Koblmuller, K. M. Sefc, N. Duftner, C. Katongo, and C. Sturmbauer. 2007. Evolutionary history of the endemic Lake Tanganyika cichlid fish Tylochromis polylepis: a recent intruder to a mature adaptive radiation. Journal Of Zoological Systematics And Evolutionary Research 45:64-71.

Laerdal, T., and M. R. Talbot. 2002. Basin neotectonics of Lakes Edward and George, East African Rift. Palaeogeography Palaeoclimatology Palaeoecology 187:213-232.

Lamb, H. F., C. R. Bates, P. V. Coombes, M. H. Marshall, M. Umer, S. J. Davies, and E. Dejen. 2007. Late Pleistocene desiccation of Lake Tana, source of the Blue Nile. Quaternary Science Reviews 26:287-299.

McCune, A. R. 1997. How fast is speciation: molecular, geological and phylogenetic evidence from adaptive radiations of fishes, Pages 585-610 *in* T. Givnish, and K. Sytsma, eds. Molecular Evolution and Adaptive Radiation, Cambridge University Press.

McCune, A. R., and N. R. Lovejoy. 1998. The relative rate of sympatric and allopatric speciation in fishes: tests using DNA sequence divergence between sister species among clades, Pages 172-185 *in* D. Howard, and S. Berloccher, eds. Endless Forms: Species and Speciation, Oxford University Press.

Menard, S. 2000. Coefficients of determination for multiple logistic regression analysis. American Statistician 54:17-24.

Moernaut, J., D. Verschuren, F. Charlet, I. Kristen, M. Fagot, and M. De Batist. 2010. The seismic-stratigraphic record of lake-level fluctuations in Lake Challa: Hydrological stability and change in equatorial East Africa over the last 140 kyr. Earth and Planetary Science Letters 290:214-223.

Owen, R. B., J. W. Barthelme, R. W. Renaut, and A. Vincens. 1982. Paleolimnology and archaeology of Holocene deposits northeast of Lake Turkana, Kenya. Nature 298:523-529.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289-290.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Experimental design and data analysis for biologists:i-xvii, 1-537.

Russell, J. M., D. Verschuren, and H. Eggermont. 2007. Spatial complexity of 'Little Ice Age' climate in East Africa: sedimentary records from two crater lake basins in western Uganda. Holocene 17:183-193.

Sato, A., N. Takezaki, H. Tichy, F. Figueroa, W. E. Mayer, and J. Klein. 2003. Origin and speciation of haplochromine fishes in east african crater lakes investigated by the analysis of their mtDNA, Mhc genes, and SINEs. Molecular Biology and Evolution 20:1448-1462.

Schliewen, U., K. Rassmann, M. Markmann, J. Markert, T. Kocher, and D. Tautz. 2001. Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. Molecular Ecology 10:1471-1488.

Schliewen, U. K., D. Tautz, and S. Paabo. 1994. Sympatric Speciation Suggested By Monophyly Of Crater Lake Cichlids. Nature 368:629-632.

Scholz, C. A., T. C. Johnson, A. S. Cohen, J. W. King, J. A. Peck, J. T. Overpeck, M. R. Talbot et al. 2007. East African megadroughts between 135 and 75 thousand years ago and bearing on early-modern human origins. Proceedings Of The National Academy Of Sciences Of The United States Of America 104:16416-16421.

Stager, J. C., and T. C. Johnson. 2008. The late Pleistocene desiccation of Lake Victoria and the origin of its endemic biota. Hydrobiologia 596:5-16.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Systematic Biology 57:758-771.

Street, F. A., and A. T. Grove. 1979. GLOBAL MAPS OF LAKE-LEVEL FLUCTUATIONS SINCE 30,000 YR BP. Quaternary Research 12:83-118.

Street-Perrott, F. A., and R. A. Perrott. 1993. Holocene vegetation, lake levels, and climate of Africa. Global climates since the last glacial maximum:318-356.

Supplementary Table 1. Species used in full phylogenetic analysis, and GenBank numbers of genes associated with these species. Yellow highlighted boxes are new sequences generated for this study.

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Alcolapia alcalicus	GQ167781	GQ167970	15428641	18072278	GQ168284	GQ168033	GQ168095	GQ168221	GQ168158
Alcolapia grahami	—	—	18072409	18072346	—	—	—	—	—
Alcolapia latilabris	—	—	18072788	18072760	—	—	—	—	—
Alcolapia ndalalani	—	—	18072780	18072748	—	—	—	—	—
Alticorpus mentale	—	—	62637864	—	—	—	—	—	—
Alticorpus pectinatum	AF305287	—	22531769	—	—	—	—	—	—
Altolamprologus calvus	DQ055011	—	509409	509407	—	—	—	—	—
Altolamprologus compressiceps	DQ055022	18182226	509408	18265830	—	—	—	—	—
Aristochromis christyi	EF585282	18182227	—	—	—	—	—	—	—
Asprotilapia leptura	AY337772.1	—	313039	313037	—	—	—	—	—
Astatoreochromis alluaudi	EU753923.1	7576476	7595677	18265818	—	—	—	—	U70339
Astatotilapia aeneocolorLA	x	—	—	—	—	—	—	—	—
Astatotilapia bloyetiECR	x	—	2394105	2394123	—	—	—	—	—
Astatotilapia brownae	—	—	—	2394141	—	—	—	—	—
Astatotilapia burtoniLTR	x	—	30143252	313030	—	—	—	—	—
Astatotilapia calliptera_KS	EU753934.1	—	—	EU753883	—	—	—	—	—
Astatotilapia calliptera_LM	AY930090	134268647	22531760	134303313	—	—	—	—	EF470867
Astatotilapia calliptera_LZ	x	—	—	—	—	—	—	—	—
Astatotilapia desfontainiiTUN	x	—	—	—	—	—	—	—	—
AstatotilapiaelegansLE	x	—	—	—	—	—	—	—	—
AstatotilapiaflavijosephiIS	x	—	47057416	—	—	—	—	—	—
AstatotilapiamacropsoidesLE	x	—	—	—	—	—	—	—	—
Astatotilapia nubila	AF305241	—	22531723	—	—	—	—	—	—
Astatotilapia orange shoulderLE	x	—	—	—	—	—	—	—	—
Astatotilapia paludinos	AY930107	—	60550047	—	—	—	—	—	—
Astatotilapia piceata	—	18182248	48773071	—	—	—	—	—	—
Astatotilapia red chestLE	x	—	—	—	—	—	—	—	—
Astatotilapia sparsidens	—	—	2394110	2394137	—	—	—	—	—
Astatotilapia stappersii	AY930046.1	—	60549994	—	—	—	—	—	—
AstatotilapiatweddleiLC	x	—	x	—	—	—	—	—	—
Astatotilapia velifer	—	—	7595614	2394121	—	—	—	—	—
Aulonocara baenschi	—	18182228	—	—	—	—	—	—	—
Aulonocranus dewindti	AY337782.1	—	33555534	313033	—	—	—	—	—
Aulonocara jacobfreibergi	—	—	11602481	—	—	—	—	—	—
Aulonocara stuartgranti	EU661720	—	—	—	—	—	—	—	—
Baileychromis centropomoides	AY682509	—	55275999	—	—	—	—	—	—
Bathybates fasciatus	AY663732	—	52221335	52221306	—	—	—	—	—
Bathybates ferox	AY663736	GQ168020	52221339	313040	GQ168335	GQ168083	GQ168146	GQ168272	GQ168209
Bathybates graueri	AY663723	—	52221327	52221290	—	—	—	—	—
Bathybates hornii	AY663735.1	—	52221338	52221312	—	—	—	—	—
Bathybates leo	AY663729	—	52221331	52221300	—	—	—	—	—
Bathybates minor	AY663721	—	52221323	52221284	—	—	—	—	—
Bathybates vittatus	AY663727	—	52221329	52221296	—	—	—	—	—
Benthochromis melanoides	AY682512.1	—	55276001	—	—	—	—	—	—
Benthochromis sp.	—	—	—	—	GQ168337	GQ168085	GQ168148	GQ168274	GQ168211
Benthochromis tricoti	AF317264	—	18029972	18265832	—	—	—	—	—
Boulengerochromis microlepis	EF679235	5114128	509464	509463	GQ168323	GQ168071	GQ168134	GQ168260	GQ168197
Buccochromis atritaeniatus	—	—	13235053	—	—	—	—	—	—
Buccochromis heterotaenia	EU661719.1	—	—	—	—	—	—	—	—
Buccochromis lepturus	U07241	—	—	—	—	—	—	—	—
Buccochromis nototaenia	—	—	62637867	—	—	—	—	—	—
Buccochromis oculatus	AF305300	—	22531782	—	—	—	—	—	—
Callochromis macrops	U07242	—	33355580	33356048	—	—	—	—	—
Callochromis melanostigma	—	—	33355576	33355998	—	—	—	—	—
Callochromis pleurospilus	AY337771.1	18182229	313055	313053	—	—	—	—	—
Callochromis stappersii	AY337775	—	33355578	33356018	—	—	—	—	—
Cardiopharynx schoutedeni	AY337791.1	—	33355531	313056	—	—	—	—	—
Chalinochromis brichardi	EF679241	5114130	509739	509737	—	—	—	—	—
Chalinochromis popeleni	U07244	33090454	—	—	—	—	—	—	—
Champsochromis spilorhynchus	U07245	—	529343	—	—	—	—	—	—
Cheilochromis euchilus	AY930092.1	—	60550032	—	—	—	—	—	—
Chetia brevicauda	EU753924.1	—	58866308	—	—	—	—	—	—
Chetia flaviventris	EU753926.1	—	58866303	—	—	—	—	—	—
Chilochromis duponti	GQ167776	GQ167965	—	—	GQ168279	GQ168028	GQ168090	GQ168216	GQ168153
Chilotilapia rhoadesii	—	—	62637872	—	—	—	—	—	—
Chromidotilapia guntheri	AF317270.1	—	—	—	—	—	—	—	—
Copadichromis borleyi	AF305308	—	22531790	—	—	—	—	—	—
Copadichromis chrysonotus	—	—	62637869	—	—	—	—	—	—
Copadichromis conophoros	—	—	541597	—	—	—	—	—	—
Copadichromis cyclicos	—	—	541595	—	—	—	—	—	—
Copadichromis eucinostomus	—	—	541593	—	—	—	—	—	—
Copadichromis mbenjii	EF585255	—	—	—	—	—	—	—	—
Copadichromis prostoma	EU661715	—	—	—	—	—	—	—	—
Copadichromis quadrimaculatus	AF305310	—	22531792	—	—	—	—	—	—
Copadichromis thinos	—	—	541602	—	—	—	—	—	—
Copadichromis virginalis	AF305281	—	22531764	—	—	—	—	—	—
Corematodus taeniatus	—	—	62637874	—	—	—	—	—	—
Ctenochromis horei	AY930100.1	—	34495295	38202260	—	—	—	—	—
Ctenochromis pectoralis	EU753938.1	—	—	2394143	—	—	—	—	—
Ctenochromis polli	EU753941.1	—	—	—	—	—	—	—	—
Ctenopharynx intermedius	—	—	62637876	—	—	—	—	—	—
Ctenopharynx pictus	EF585254	—	—	—	—	—	—	—	—
Cunningtonia longiventralis	AY682516.1	—	33355529	313049	—	—	—	—	—
Cyathochromis obliquidens	—	—	1881618	—	—	—	—	—	—
Cyathopharynx foae	—	—	50262139	—	—	—	—	—	—
Cyathopharynx furcifer	AY337781.1	—	50262143	313043	—	—	—	—	—

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Cyclopharynx fwae	AY930099.1	---	18029958	18265820	---	---	---	---	---
Cymatogaster aggregata	x	AY662711	---	AF370623	---	---	---	---	AY662811
Cynotilapia afra	EF585264	---	58866394	---	---	---	---	---	---
Cyphotilapia frontosa	EF679242	---	393082	18265842	---	---	---	---	---
Cyprichromis cf. leptosoma 'yellow'	---	---	---	58379197	---	---	---	---	---
Cyprichromis leptosoma	AY337786	GQ168023	58379152	18265812	GQ168338	GQ168086	GQ168149	GQ168275	GQ168212
Cyprichromis microlepidotus	AY740354.1	---	58379159	58379279	---	---	---	---	---
Cyprichromis pavo	AY740382	---	58379186	58379285	---	---	---	---	---
Cyprichromis sp. 'jumbo'	---	---	---	18265810	---	---	---	---	---
Cyprichromis sp. 'zebra'	---	---	---	58379283	---	---	---	---	---
Cyprichromis zonatus	AY740377.1	---	58379187	58379337	---	---	---	---	---
Cyrtocara moorii	AY930089.1	18182230	529344	---	---	---	---	---	---
Dimidiochromis compressiceps	EF585267	---	13235093	---	---	---	---	---	---
Dimidiochromis kiwinge	AF305322	---	22531804	---	---	---	---	---	---
Dimidiochromis strigatus	---	18182231	---	---	---	---	---	---	---
Diplotaxodon aeneus	---	---	62637885	---	---	---	---	---	---
Diplotaxodon apogon	---	---	62637887	---	---	---	---	---	---
Diplotaxodon argenteus	---	---	62637889	---	---	---	---	---	---
Diplotaxodon brevimaxillaris	AF305264	---	22531746	---	---	---	---	---	---
Diplotaxodon greenwoodi	AF305269.1	134268645	22531752	134303321	---	---	---	---	EF470868
Diplotaxodon holochromis	AF305262.1	---	22531744	---	---	---	---	---	---
Diplotaxodon limnothrissa	AF305261	---	22531738	---	---	---	---	---	---
Diplotaxodon limnothrissa black pelvic	---	---	116178671	---	---	---	---	---	---
Diplotaxodon macrops	AF305266	---	62637903	---	---	---	---	---	---
Diplotaxodon macrops offshore	---	---	116178708	---	---	---	---	---	---
Diplotaxodon similis	AF305271	---	22531756	---	---	---	---	---	---
Docimodus evelynae	EF585252	---	---	---	---	---	---	---	---
Eclectochromis ornatus	EU661717.1	---	---	---	---	---	---	---	---
Ectodus descampsii	AY337790.1	18182232	18029950	313060	---	---	---	---	---
Embiotoca jacksoni	---	AY662712	---	AF159331	---	---	---	---	AY662812
Enantiopus melanogenys	AY682517	---	33355554	33356000	---	---	---	---	---
Enterochromis cinctus	---	---	7595588	---	---	---	---	---	---
Enterochromis spLE	x	---	---	---	---	---	---	---	---
Eretmodus cyanostictus	AF398220	GQ168019	5918049	5918033	GQ168334	GQ168082	GQ168145	GQ168271	GQ168208
Etia nguti	GQ167777	58199021	---	---	GQ168280	GQ168029	GQ168091	GQ168217	GQ168154
Etroplus canarensis	---	AY662713	---	---	---	---	---	---	AY662816
Etroplus maculatus	x	EF095604	---	---	---	---	---	---	AY662818
Etroplus suratensis	---	AY263829	---	---	---	---	---	---	AY662817
Exochochromis anagenys	---	---	22531797	---	---	---	---	---	---
Fossorochromis rostratus	EF585281	---	62637908	---	---	---	---	---	---
Gaurochromis simpsoni	---	33090458	7595580	---	---	---	---	---	---
Gaurochromis spLE	x	---	---	---	---	---	---	---	---
Genyochromis mento	AF305297	---	1881619	---	---	---	---	---	---
Gephyrochromis lawsi	---	---	62637910	---	---	---	---	---	---
Gnathochromis permaxillaris	AY682519	18182233	55276010	18265834	---	---	---	---	---
Gnathochromis pfefferi	EF679245	---	18029974	18265836	---	---	---	---	---
Gobiocichla ethelwynnae	---	58199022	---	---	---	---	---	---	---
Gobiocichla wonderi	GQ167778	GQ167967	---	---	GQ168281	GQ168030	GQ168092	GQ168218	GQ168155
Grammatotria lemairii	AY337787.1	18182234	313068	313063	---	---	---	---	---
Greenwoodochromis bellcrossi	AY682523	---	55276012	---	---	---	---	---	---
Greenwoodochromis christyi	AY682525	---	55276016	---	---	---	---	---	---
Haplochromis adolfifrederici	---	---	30143182	---	---	---	---	---	---
Haplochromis astatodon	---	---	30143158	---	---	---	---	---	---
Haplochromis chala	---	---	2394106	2394129	---	---	---	---	---
Haplochromis crebridens	---	---	30143181	---	---	---	---	---	---
Haplochromis gracilior	AY930079	---	30143255	---	---	---	---	---	---
Haplochromis graueri	---	---	30143123	---	---	---	---	---	---
Haplochromis insidiae	AY930077.1	---	30143094	---	---	---	---	---	---
Haplochromis lividus	---	---	7595585	---	---	---	---	---	---
Haplochromis microchrysomelas	---	---	30143162	---	---	---	---	---	---
Haplochromis nigroides	---	---	30143144	---	---	---	---	---	---
Haplochromis obliquidens	AY930097	---	60550037	---	---	---	---	---	---
Haplochromis occultidens	---	---	30143133	---	---	---	---	---	---
Haplochromis olivaceus	---	---	30143126	---	---	---	---	---	---
Haplochromis paucidens	---	---	30143108	---	---	---	---	---	---
Haplochromis purple yellowLV	x	---	---	---	---	---	---	---	---
Haplochromis rubescens	---	---	30143088	---	---	---	---	---	---
Haplochromis scheffersi	---	---	30143097	---	---	---	---	---	---
Haplochromis sp. crebridens/olivaceus	---	---	30143113	---	---	---	---	---	---
Haplochromis sp. Fayoum	EU753945.1	---	---	---	---	---	---	---	---
Haplochromis sp. Kanyaboli	EU753944.1	---	---	---	---	---	---	---	---
Haplochromis sp. Kisangani	AY930062.1	---	---	---	---	---	---	---	---
Haplochromis sp. Mburo Black	EU753946.1	---	---	---	---	---	---	---	---
Haplochromis sp. nov.	EU753928	---	---	EU753877	---	---	---	---	---
Haplochromis thick skinLV	x	---	---	---	---	---	---	---	---
Haplochromis vittatus	---	---	30143178	---	---	---	---	---	---
Haplotaxodon microlepis	EF437498.1	---	55276018	EF679278	---	---	---	---	---
Haplotaxodon trifasciatus	AY682531	---	55276020	---	---	---	---	---	---
Harpagochromis guiarti	---	---	48773073	---	---	---	---	---	---
Harpagochromis spLE	x	---	---	---	---	---	---	---	---
Harpagochromis squamipinnis	EU753943.1	---	30143214	---	---	---	---	---	---
Hemibates stenosoma	AY663719.1	---	52221320	52221276	---	---	---	---	---
Hemichromis bimaculatus	---	4091094	---	2394111	---	---	---	---	---
Hemichromis elongatus	AY663714.1	GQ168001	112735225	---	GQ168315	GQ168063	GQ168126	GQ168252	GQ168189
Hemichromis guttatus	---	58199023	---	---	---	---	---	---	AY662866
Hemitilapia oxyrhyncha	EF585277	---	---	---	---	---	---	---	---
Heterochromis multidens	GQ167779	4321412	---	18265806	GQ168282	GQ168031	GQ168093	GQ168219	AF113060
Iodotropheus sprengerae	---	---	1881620	---	---	---	---	---	---
Iranocichla hormuzensis	GQ167830	GQ168018	---	---	GQ168333	GQ168081	GQ168144	GQ168270	GQ168207
Julidochromis affinis	---	---	510122	---	---	---	---	---	---

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Julidochromis brichardi	EF462232	—	—	—	—	—	—	—	—
Julidochromis dickfeldi	EF462230	—	—	—	—	—	—	—	—
Julidochromis marlieri	DQ055039	—	510121	510120	—	—	—	—	—
Julidochromis ornatus	EF191082	—	47116853	—	—	—	—	—	—
Julidochromis regani	EF462228	134268652	393079	340559	—	—	—	—	EF470870
Julidochromis transcriptus	EF462231	—	—	—	—	—	—	—	—
Konia dikume	AJ845104	—	55468948	x	—	—	—	—	—
Konia eisentrauti	AJ845102	—	55468946	x	—	—	—	—	—
Labeotropheus fuelleborni	EF585259	—	1881622	—	—	—	—	—	—
Labeotropheus trewavasae	EF585283	—	7595685	—	—	—	—	—	—
Labidochromis caeruleus	AY740383	7576478	10046829	134303329	—	—	—	—	EF470871
Labidochromis gigas	EF585276	—	—	—	—	—	—	—	—
Labidochromis vellicans	—	—	62637911	—	—	—	—	—	—
Labrochromis ishmaeli	—	18182235	—	—	—	—	—	—	—
Lamprologus callipterus	EF462258	—	47116775	510127	—	—	—	—	—
Lamprologus congoensis	AY740385	—	510132	510128	—	—	—	—	—
Lamprologus cunningtoni	x	—	—	x	—	—	—	—	—
Lamprologus kungweensis	EF191084	—	—	—	—	—	—	—	—
Lamprologus laparogramma	EF462278	—	—	—	—	—	—	—	—
Lamprologus lemairii	EF191093	—	47116856	x	—	—	—	—	—
Lamprologus meleagris	EF191097	—	67553148	—	—	—	—	—	—
Lamprologus mocquardi	AF398225	—	510142	510141	—	—	—	—	—
Lamprologus ocellatus	EF191114	—	—	—	—	—	—	—	—
Lamprologus ornatipinnis	EF191110	—	—	—	—	—	—	—	—
Lamprologus savoryi	x	—	—	x	—	—	—	—	—
Lamprologus signatus	EF191086	—	—	—	—	—	—	—	—
Lamprologus speciosus	DQ055032	—	67553153	—	—	—	—	—	—
Lamprologus teugelsi	DQ055059	—	67553174	—	—	—	—	—	—
Lamprologus werneri	—	—	510144	510143	—	—	—	—	—
Lepidiolamprologus attenuatus	AY682532.1	—	55276021	x	—	—	—	—	—
Lepidiolamprologus cunningtoni	DQ055053	—	50916249	—	—	—	—	—	—
Lepidiolamprologus elongatus	DQ055021	18182237	67553142	510135	—	—	—	—	—
Lepidiolamprologus kendalli	DQ055042	—	67553160	—	—	—	—	—	—
Lepidiolamprologus nkambae	DQ055046	—	67553165	—	—	—	—	—	—
Lepidiolamprologus profundicola	DQ055025	—	47116841	—	—	—	—	—	—
Lepidolamprologus elongatus	EF679248	—	—	EF679280	—	—	—	—	—
Lestradea perspicax	AY337765	—	313074	313072	—	—	—	—	—
Lestradea stappersii	AY337792.1	—	33355527	33356028	—	—	—	—	—
Lethrinops altus	—	—	62637918	—	—	—	—	—	—
Lethrinops auritus	U07252	—	529341	—	—	—	—	—	—
Lethrinops furcifer	AF305316	—	22531798	—	—	—	—	—	—
Lethrinops gossei	AF305290	—	22531772	—	—	—	—	—	—
Lethrinops longipinnis	AF305295	—	22531777	—	—	—	—	—	—
Lethrinops microdon	AF305292	—	22531774	—	—	—	—	—	—
Lethrinops mylodon	—	—	62637926	—	—	—	—	—	—
Lethrinops oliveri	AF305288	—	22531770	—	—	—	—	—	—
Lethrinops polli	—	—	62637931	—	—	—	—	—	—
Limnochromis abeelei	AY682533	—	55276022	—	—	—	—	—	—
Limnochromis auritus	AF398216	—	313071	313069	—	—	—	—	—
Limnochromis staneri	AY682538	—	55276030	—	—	—	—	—	—
Limnotilapia dardennii	DQ093109	—	34495299	38202268	—	—	—	—	—
Lipochromis cryptodonLV	x	—	—	—	—	—	—	—	—
Lipochromis maxillaris	—	—	60550087	—	—	—	—	—	—
Lipochromis melanopterus	—	—	7595589	—	—	—	—	—	—
Lipochromis obesus	—	—	529342	—	—	—	—	—	—
Lipochromis velvet cryptodonLV	x	—	—	—	—	—	—	—	—
Lobochilotes labiatus	EF679250	—	393083	18265844	—	—	—	—	—
Maravichromis mola	EF585274	—	—	—	—	—	—	—	—
Maylandia callainos	EF585271	—	7595682	—	—	—	—	—	—
Maylandia estherae	—	18182238	—	—	—	—	—	—	—
Maylandia zebra	DQ093114.1	33090452	60550078	—	—	—	—	—	EF470875
Mchenga eucinostomus	EF585268	—	—	—	—	—	—	—	—
Melanochromis auratus	AY930069.1	—	403987	—	—	—	—	—	—
Melanochromis elastodema	—	—	62637935	—	—	—	—	—	—
Melanochromis heterochromis	—	—	403997	—	—	—	—	—	—
Melanochromis johannii	—	—	404001	—	—	—	—	—	—
Melanochromis melanopterus	—	—	404005	—	—	—	—	—	—
Melanochromis parallelus	—	—	404006	—	—	—	—	—	—
Melanochromis simulans	—	—	404003	—	—	—	—	—	—
Melanochromis vermivorus	EF585270	—	—	—	—	—	—	—	—
Metriaclima zebra	—	—	—	134303335	—	—	—	—	—
Microdontochromis rotundiventralis	AY337793.1	—	33355551	33356100	—	—	—	—	—
Microdontochromis tenuidentatus	AY337784.1	—	33355549	313075	—	—	—	—	—
Myaka myaka	AJ845106	—	55468950	x	—	—	—	—	—
Mylochromis anaphyrmus	AF305321	—	22531803	—	—	—	—	—	—
Mylochromis ericotaenia	—	—	62637938	—	—	—	—	—	—
Mylochromis labidodon	—	—	13235351	—	—	—	—	—	—
Mylochromis lateristriga	—	—	7595686	—	—	—	—	—	—
Nanochromis parilus	—	GQ168003	—	—	GQ168317	GQ168065	GQ168128	GQ168254	GQ168191
Neochromis rufocaudalis	—	18182245	7595607	—	—	—	—	—	—
Neolamprologus bifasciatus	EF462240	—	—	—	—	—	—	—	—
Neolamprologus boulengeri	—	—	67553155	—	—	—	—	—	—
Neolamprologus brevis	DQ055020	—	510159	—	—	—	—	—	—
Neolamprologus brichardi	EF462245	33090455	510161	510158	—	—	—	—	—
Neolamprologus buescheri	DQ055033	—	67553154	—	—	—	—	—	—
Neolamprologus calliurus	EF191083	—	510163	—	—	—	—	—	—
Neolamprologus caudopunctatus	AY740388	—	108793969	—	—	—	—	—	—
Neolamprologus christyi	AY740389.1	—	510157	510162	—	—	—	—	—
Neolamprologus cylindricus	EF462224	—	510165	—	—	—	—	—	—
Neolamprologus devosi	EF437476.1	—	—	—	—	—	—	—	—

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Neolamprologus falicula	EF462246	---	---	---	---	---	---	---	---
Neolamprologus fasciatus	EF191119	---	47116843	---	---	---	---	---	---
Neolamprologus furcifer	EF462249	---	510168	510167	---	---	---	---	---
Neolamprologus gracilis	---	---	18254012	18307833	---	---	---	---	---
Neolamprologus hecqui	DQ055041	---	67553139	---	---	---	---	---	---
Neolamprologus helianthus	DQ055013	---	18254007	18307823	---	---	---	---	---
Neolamprologus leleupi	DQ093113	---	---	---	---	---	---	---	---
Neolamprologus leloupi	EF191103.1	---	---	---	---	---	---	---	---
Neolamprologus longicaudata	EF462250	---	510171	510169	---	---	---	---	---
Neolamprologus marunguensis	AY740390.1	---	18254002	18307813	---	---	---	---	---
Neolamprologus meeli	DQ055051	---	67553169	---	---	---	---	---	---
Neolamprologus modestus	DQ055012	---	47116917	---	---	---	---	---	---
Neolamprologus mondabu	EF462241	---	47116855	x	---	---	---	---	---
Neolamprologus multifasciatus	EF191089	---	---	---	---	---	---	---	---
Neolamprologus mustax	EF462223	---	---	---	---	---	---	---	---
Neolamprologus niger	AY740391.1	---	---	---	---	---	---	---	---
Neolamprologus nigriventris	AY740392.1	---	---	---	---	---	---	---	---
Neolamprologus olivaceous	AY740393.1	---	18254004	18307817	---	---	---	---	---
Neolamprologus palmeri	AY740394.1	---	---	---	---	---	---	---	---
Neolamprologus pectoralis	EF462238	---	---	---	---	---	---	---	---
Neolamprologus petricola	---	---	47116831	---	---	---	---	---	---
Neolamprologus prochilus	EF462248	---	---	---	---	---	---	---	---
Neolamprologus pulcher	EF462244	---	47116778	18307839	---	---	---	---	---
Neolamprologus savoryi	EF462247	---	18254001	18307807	---	---	---	---	---
Neolamprologus similis	EF191100	---	67553151	---	---	---	---	---	---
Neolamprologus splendens	---	---	18254013	18307835	---	---	---	---	---
Neolamprologus tetracanthus	EF462220	---	50916263	---	---	---	---	---	---
Neolamprologus toae	x	---	510176	510175	---	---	---	---	---
Neolamprologus tretocephalus	EF462219	---	47116857	EF679285	---	---	---	---	---
Neolamprologus variostigma	DQ055028	---	67553149	---	---	---	---	---	---
Neolamprologus ventralis	EF462233	---	---	---	---	---	---	---	---
Neolamprologus wauthioni	EF191116	---	---	---	---	---	---	---	---
Nimbochromis fuscotaeniatus	---	134268656	---	134303333	---	---	---	---	EF470872
Nimbochromis linni	EF585279	---	58866395	---	---	---	---	---	---
Nimbochromis livingstonii	EU753948.1	---	62637942	---	---	---	---	---	---
Nimbochromis polystigma	EF585262	---	13235349	---	---	---	---	---	---
Nimbochromis venustus	EU753947.1	---	11602783	---	---	---	---	---	---
Ophthalmotilapia boops	AY337773.1	---	50262165	33356010	---	---	---	---	---
Ophthalmotilapia heterodonta	---	---	12830449	---	---	---	---	---	---
Ophthalmotilapia nasuta	AY337783.1	---	12830454	33356070	---	---	---	---	---
Ophthalmotilapia ventralis	AY337774.1	---	33355525	313081	---	---	---	---	---
Opthalmotilapia heterodonta	EF679254	---	---	EF679286	---	---	---	---	---
Oreochromis amphimelas	AF317230.1	---	15428675	18076051	---	---	---	---	---
Oreochromis andersonii	GQ167805	GQ167994	15428673	---	GQ168308	GQ168056	GQ168119	GQ168245	GQ168182
Oreochromis aureus	DQ465029	90018769	24635221	14161580	---	---	---	---	---
Oreochromis esculentus	AF317232.1	58199024	15428667	18076057	---	---	---	---	---
Oreochromis jipe	---	---	15428678	---	---	---	---	---	---
Oreochromis karongae	DQ465030	134268654	15429056	134303325	---	---	---	---	EF470873
Oreochromis leucostictus	AF317233.1	---	116672807	---	---	---	---	---	---
Oreochromis macrochir	AF317235.1	---	---	---	---	---	---	---	---
Oreochromis malagarasi	---	---	15428649	2394115	---	---	---	---	---
Oreochromis mortimeri	---	---	15429057	---	---	---	---	---	---
Oreochromis mossambicus	AF317234.1	33090451	15428652	4903283	---	---	---	---	---
Oreochromis mweruensis	AF317236	---	---	---	---	---	---	---	---
Oreochromis niloticus	AF317237	5114129	15429060	24635169	GQ168283	GQ168032	GQ168094	GQ168220	GQ168157
Oreochromis niloticus baringoensis	---	---	---	18076054	---	---	---	---	---
Oreochromis schwebischi	AF317238.1	---	---	---	---	---	---	---	---
Oreochromis tanganicae	AF317240	GQ167971	313080	13224	GQ168285	GQ168034	GQ168096	GQ168222	GQ168159
Oreochromis urolepis	AF317239.1	---	15428653	---	---	---	---	---	---
Oreochromis variabilis	AF317241.1	---	---	---	---	---	---	---	---
Orthochromis cf. kalungwishiensis	---	---	x	---	---	---	---	---	---
Orthochromis kalungwishiensis	---	---	x	---	---	---	---	---	---
Orthochromis kasuluensis	AY930049.1	---	60549997	---	---	---	---	---	---
Orthochromis luichensis	AY930052.1	---	60550000	---	---	---	---	---	---
Orthochromis malagaraziensis	AY930054.1	---	60550002	2394145	---	---	---	---	---
Orthochromis mazimeroensis	AY930053.1	---	60550001	18265828	---	---	---	---	---
Orthochromis mosoensis	AY930055.1	---	60550003	---	---	---	---	---	---
Orthochromis polyacanthus	AF398231	---	18029959	18265822	---	---	---	---	---
Orthochromis rubrolabialis	AY930051.1	---	60549999	---	---	---	---	---	---
Orthochromis rugufuensis	AY930050.1	---	60549998	---	---	---	---	---	---
Orthochromis uvinzae	AY930048.1	---	60549996	---	---	---	---	---	---
Otopharynx argyrosoma	---	---	62637944	---	---	---	---	---	---
Otopharynx brooksi	AF305303	---	13235094	---	---	---	---	---	---
Otopharynx heterodon	EF585278	---	---	---	---	---	---	---	---
Otopharynx speciosus	AF305304	---	22531786	---	---	---	---	---	---
Otopharynx walteri	EU661716.1	---	---	---	---	---	---	---	---
Oxylapia polli	AF317275.1	AY263817	---	---	---	---	---	---	AY662832
Pallidochromis tokolosh	AF305276	---	22531758	---	---	---	---	---	---
Paracyprichromis brieni	AF398223	---	313089	313087	---	---	---	---	---
Paracyprichromis nigripinnis	AY740339.1	---	58379145	58379257	---	---	---	---	---
Paralabidochromis beadlei	---	---	7595581	---	---	---	---	---	---
Paralabidochromis chilotes	---	---	7595587	---	---	---	---	---	---
Paralabidochromis plagiodon	---	---	7595591	---	---	---	---	---	---
Paralabidochromis rockkribensisLV	x	---	---	---	---	---	---	---	---
Paratilapia polleni	x	AY263818	---	---	---	---	---	---	---
Paratilapia polleni Nosy Be	---	AY662719	---	---	---	---	---	---	AY662834
Paratilapia polleni Ravelobe	---	AY662720	---	---	---	---	---	---	AY662835
Paretroplus dambabe	---	AY263822	---	---	---	---	---	---	AY662819
Paretroplus kieneri	---	AY263827	---	---	---	---	---	---	AY662821
Paretroplus maculatus	x	AY263820	---	---	---	---	---	---	AY662824

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Paretroplus maromandia	--	AY263821	--	--	--	--	--	--	AY662825
Paretroplus menarambo	--	AY263823	--	--	--	--	--	--	AY662826
Paretroplus nourissati	--	AY263828	--	--	--	--	--	--	AY662827
Paretroplus polyactis	--	AF112582	--	--	--	--	--	--	U70327
Paretroplus polyactis North	--	AY662718	--	--	--	--	--	--	AY662831
Paretroplus polyactis South	--	AY263826	--	--	--	--	--	--	AY662828
Paretroplus tsimoly	--	AY662716	--	--	--	--	--	--	AY662829
Pelmatochromis buettikoferi	GQ167783	GQ167972	--	--	GQ168286	GQ168035	GQ168097	GQ168223	GQ168160
Pelmatochromis nigrofasciatus	--	58199025	--	--	GQ168287	GQ168036	GQ168098	GQ168224	AY662870
Pelvicachromis humilis	--	--	2394099	2394113	--	--	--	--	--
Pelvicachromis pulcher	AF317271.1	58199026	--	2394149	--	--	--	--	EF470874
Perissodus eccentricus	EF437506	--	--	--	--	--	--	--	--
Perissodus microlepis	EF437483.1	--	18029977	18265838	--	--	--	--	--
Petrochromis ephippium	--	--	47156788	38202274	--	--	--	--	--
Petrochromis famula	EF679265	--	34495303	38202276	--	--	--	--	--
Petrochromis fasciolatus	--	--	50916259	--	--	--	--	--	--
Petrochromis kazumbe	x	--	--	x	--	--	--	--	--
Petrochromis kazumbe "polyodon"	x	--	--	x	--	--	--	--	--
Petrochromis macrognathus	AY930068.1	--	50916247	--	--	--	--	--	--
Petrochromis moshi	x	--	--	x	--	--	--	--	--
Petrochromis orthognathus	x	--	393084	18265846	--	--	--	--	--
Petrochromis polyodon	x	--	50916264	x	--	--	--	--	--
Petrochromsi sp. "moshi"	x	--	--	x	--	--	--	--	--
Petrochromis trewavasae	--	--	47156790	38202300	--	--	--	--	--
Petrotilapia nigra	EU661721.1	--	--	--	--	--	--	--	--
Pharyngochromis acuticeps	AY930094	--	58866300	--	--	--	--	--	--
Pharyngochromis sp.	--	--	x	--	--	--	--	--	--
Pharyngochromis sp. DAJ-2005	--	--	58866322	--	--	--	--	--	--
Placidochromis cf. subocularis MRI-2005	--	--	62637950	--	--	--	--	--	--
Placidochromis johnstoni	EF585269	--	--	--	--	--	--	--	--
Placidochromis milomo	EF585251	--	393087	--	--	--	--	--	--
Platytaeniodus degeni	AY930064.1	18182249	48773064	--	--	--	--	--	--
Plecodus elaviae	EF437504	--	--	--	--	--	--	--	--
Plecodus multidentatus	EF437505.1	--	--	--	--	--	--	--	--
Plecodus paradoxus	EF437499	--	--	--	--	--	--	--	--
Plecodus straeleni	AF398221	--	313092	313090	--	--	--	--	--
Prognathochromis dentex	--	--	48773070	--	--	--	--	--	--
Prognathochromis longirostris	--	--	48773069	--	--	--	--	--	--
Prognathochromis paraguiarti	--	--	48773072	--	--	--	--	--	--
Prognathochromis spLE	x	--	--	--	--	--	--	--	--
Prognathochromis venator	--	--	7595599	--	--	--	--	--	--
Protomelas annectens	EU661718.1	--	13235352	--	--	--	--	--	--
Protomelas fenestratus	AF305301	--	7595687	--	--	--	--	--	--
Protomelas insignis	--	--	62637951	--	--	--	--	--	--
Protomelas similis	EU661714.1	--	--	--	--	--	--	--	--
Protomelas spilopterus	EF585253	--	--	--	--	--	--	--	--
Protomelas taeniolatus	AF305302	--	22531784	--	--	--	--	--	--
Psammochromis riponianus	--	--	7595592	--	--	--	--	--	--
Pseudocrenilabrus broad head black pelvic	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus dwarf black pelvic	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus fire tail	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus green weed picker	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus grey back	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus grey moeruensis	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus long brown	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus long grey	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus machadoi	EU753936.1	--	58866315	--	--	--	--	--	--
Pseudocrenilabrus multicolor	AY930070.1	--	18029960	18265824	--	--	--	--	--
Pseudocrenilabrus multicolor victoriae	AY930070	--	60550018	--	--	--	--	--	--
Pseudocrenilabrus nicholsi	AY602994	--	47498977	47118406	--	--	--	--	--
Pseudocrenilabrus pale deep	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus philander	AY602993.1	--	47717261	47118404	--	--	--	--	--
Pseudocrenilabrus sp. blue Lunzua River	EU753951.1	--	47717289	--	--	--	--	--	--
Pseudocrenilabrus sp. orange Mwatishi River	EU753952.1	--	47717269	--	--	--	--	--	--
Pseudocrenilabrus telmatochromis like	--	--	x	--	--	--	--	--	--
Pseudocrenilabrus weed picker	--	--	x	--	--	--	--	--	--
Pseudosimochromis curvifrons	--	--	18029982	18265848	--	--	--	--	--
Pseudotropheus aurora	EF585266	--	--	--	--	--	--	--	--
Pseudotropheus barlowi	--	--	1881627	--	--	--	--	--	--
Pseudotropheus crabro	EF585256	--	--	--	--	--	--	--	--
Pseudotropheus elongatus	EF585272	--	--	--	--	--	--	--	--
Pseudotropheus lanisticola	--	--	62637956	--	--	--	--	--	--
Pseudotropheus livingstonii	AY930061.1	--	60550009	--	--	--	--	--	--
Pseudotropheus microstoma	EF585258	--	--	--	--	--	--	--	--
Pseudotropheus tropheops	AY740384.1	--	62637959	--	--	--	--	--	--
Pseudotropheus tropheops gracilior	EF585260	--	--	--	--	--	--	--	--
Pseudotropheus williamsi	--	--	1881637	--	--	--	--	--	--
Pseudotropheus xanstomachus	--	--	1881639	--	--	--	--	--	--
Pterochromis congicus	GQ167807	GQ167974	--	--	GQ168288	GQ168037	GQ168099	GQ168225	GQ168162
Ptychochromoides betsileanus	--	AY263815	--	--	--	--	--	--	AY662838
Ptychochromis grandidieri	--	AY263811	--	--	--	--	--	--	AY662841
Ptychochromoides katria	x	--	x	--	--	--	--	--	AY662840
Ptychochromis sp. garaka	--	AY662723	--	--	--	--	--	--	AY662845
Ptychochromis oligacanthus	--	AY279667	--	--	--	--	--	--	AY279770
Ptychochromis inornatus	--	AY263812	--	--	--	--	--	--	--
Ptychochromoides sp. Makira	--	AY662724	--	--	--	--	--	--	--
Ptychochromis sp. sofia	--	AY662725	--	--	--	--	--	--	AY662847
Ptychochromoides vondrozo	--	AY263816	--	--	--	--	--	--	AY662839
Ptyochromis ishmaeli	--	--	48773062	--	--	--	--	--	--
Ptyochromis sauvagei	AY930059	--	7595593	--	--	--	--	--	--

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Ptyochromis xenognathusLV	x	—	7595594	—	—	—	—	—	—
Pungu maclareni	AJ845101.1	—	55468944	x	—	—	—	—	—
Reganochromis calliurus	AY682544	—	55276032	—	—	—	—	—	—
Rhamphochromis esox	AF305252.1	134268657	22531734	134303331	—	—	—	—	EF470876
Rhamphochromis leptosoma	AF305253.1	—	62637967	—	—	—	—	—	—
Rhamphochromis longicepsLM	x	—	62637971	—	—	—	—	—	—
Rhamphochromis macrophthalmus	AF305249	—	22531731	—	—	—	—	—	—
Rhamphochromis sp. big mouth	AF305251.1	—	22531733	—	—	—	—	—	—
Rhamphochromis sp. brown	AF305247.1	—	22531729	—	—	—	—	—	—
Rhamphochromis sp. long snout	—	—	62637973	—	—	—	—	—	—
Rhamphochromis sp. maldeco	AF305254.1	—	62637975	—	—	—	—	—	—
Sargochromis big miller	—	—	x	—	—	—	—	—	—
Sargochromis carlottae	EF393682	—	58866304	—	—	—	—	—	—
Sargochromis cf. mellandi	—	—	x	—	—	—	—	—	—
Sargochromis cf. mortimeri DAJ-2005	—	—	58866327	—	—	—	—	—	—
Sargochromis codringtonii	EF393713	—	58866306	—	—	—	—	—	—
Sargochromis compressed suborbital	—	—	x	—	—	—	—	—	—
Sargochromis coulteri	EU753954.1	—	58866341	—	—	—	—	—	—
Sargochromis deep body yellowish	—	—	x	—	—	—	—	—	—
Sargochromis deep short jaws	—	—	x	—	—	—	—	—	—
Sargochromis elongate	—	—	x	—	—	—	—	—	—
Sargochromis giardi	AY930098.1	—	58866319	—	—	—	—	—	—
Sargochromis mellandi	EF393700	—	58866348	—	—	—	—	—	—
Sargochromis mortimeri	—	—	58866342	—	—	—	—	—	—
Sargochromis red face	—	—	x	—	—	—	—	—	—
Sargochromis sp. 1 DAJ-2005	—	—	58866324	—	—	—	—	—	—
Sargochromis sp. 2 DAJ-2005	—	—	58866340	—	—	—	—	—	—
Sargochromis sp. 3 DAJ-2005	—	—	58866325	—	—	—	—	—	—
Sargochromis sp. kafuensis	—	—	126015733	—	—	—	—	—	—
Sargochromis sp. Lisikili	—	—	x	—	—	—	—	—	—
Sargochromis sp. SK-2007	—	—	126015743	—	—	—	—	—	—
Sargochromis sp. Zambesi	—	—	x	—	—	—	—	—	—
Sargochromis thin face green	—	—	x	—	—	—	—	—	—
Sargochromis yellow face mellandi	—	—	x	—	—	—	—	—	—
Sarotherodon caroli	AJ845113	—	55468956	x	—	—	—	—	—
Sarotherodon caudomarginatus	AF317243.1	GQ167975	—	—	GQ168289	GQ168038	GQ168100	GQ168226	GQ168163
Sarotherodon galilaeus	AF317244	—	15429064	14161582	—	—	—	—	—
Sarotherodon galilaeus Cross	—	—	x	x	—	—	—	—	—
Sarotherodon galilaeus Ejagham	—	—	x	x	—	—	—	—	—
Sarotherodon galilaeus Meme	—	—	x	x	—	—	—	—	—
Sarotherodon galilaeus multifasciatus	AJ845087.1	—	55468930	—	—	—	—	—	—
Sarotherodon galilaeus sanagaensis	AJ845085.1	—	55468929	x	—	—	—	—	—
Sarotherodon linnellii	AJ845114	—	55468958	x	—	—	—	—	—
Sarotherodon lohbergeri	AJ845108	—	55468952	x	—	—	—	—	—
Sarotherodon melanotheron	AF317245.1	GQ167976	19569083	14134134	—	—	—	—	—
Sarotherodon mvogoi	GQ167811	GQ168000	—	—	GQ168314	GQ168062	GQ168125	GQ168251	GQ168188
Sarotherodon nigripinnis	AJ845084.1	—	19568978	—	GQ168290	GQ168039	GQ168101	GQ168227	GQ168164
Sarotherodon occidentalis	AF317246.1	—	—	—	—	—	—	—	—
Sarotherodon sp. aff. galilaeus mudfeeder	—	GQ167977	—	—	GQ168291	GQ168040	GQ168102	GQ168228	GQ168165
Sarotherodon sp. bighead	AJ845091.1	—	55468935	—	—	—	—	—	—
Sarotherodon sp. mudfeeder	AJ845092.1	—	55468936	—	—	—	—	—	—
Sarotherodon steinbachi	AJ845110	18182244	55468954	x	—	—	—	—	—
Schwetzochromis neodon	EU753957.1	—	—	—	—	—	—	—	—
Schwetzochromis stormsi	AY930057	—	60550005	—	—	—	—	—	—
Sciaenochromis benthicola	AF305298	—	22531780	—	—	—	—	—	—
Sciaenochromis gracilis	—	—	13235350	—	—	—	—	—	—
Sciaenochromis psammophilus	AF305324	—	22531806	—	—	—	—	—	—
Sciaenochromis spilostichus	—	—	62637940	—	—	—	—	—	—
Serranochromis altus	EF393697.1	—	58866372	—	—	—	—	—	—
Serranochromis angusticeps	EF393685	—	58866380	—	—	—	—	—	—
Serranochromis angusticeps yellow	—	—	x	—	—	—	—	—	—
Serranochromis cf. altus	—	—	x	—	—	—	—	—	—
Serranochromis cf. macrocephalus 1	—	—	x	—	—	—	—	—	—
Serranochromis checkerboard	—	—	x	—	—	—	—	—	—
Serranochromis dark long body	—	—	x	—	—	—	—	—	—
Serranochromis deep red	—	—	x	—	—	—	—	—	—
Serranochromis diplotaxodon face	—	—	x	—	—	—	—	—	—
Serranochromis longimanus	—	—	126015680	—	—	—	—	—	—
Serranochromis long body	—	—	x	—	—	—	—	—	—
Serranochromis long face blue	—	—	x	—	—	—	—	—	—
Serranochromis long pelvic	—	—	x	—	—	—	—	—	—
Serranochromis macrocephalus	EF393689	—	58866317	—	—	—	—	—	—
Serranochromis macrocephalus deep body	—	—	x	—	—	—	—	—	—
Serranochromis meridianus	—	—	58866343	—	—	—	—	—	—
Serranochromis robustus	EF393686	5114131	126015708	—	—	—	—	—	—
Serranochromis silver long body	—	—	x	—	—	—	—	—	—
Serranochromis stappersi	EF393698	—	58866385	—	—	—	—	—	—
Serranochromis thumbergi	EF393703	—	58866298	—	—	—	—	—	—
Simochromis babaulti	DQ093110	—	1110522	13504	—	—	—	—	—
Simochromis diagramma	AY930087.1	—	1055357	38202310	—	—	—	—	—
Simochromis marginatus	AY930088.1	—	1245389	x	—	—	—	—	—
Spathodus erythrodon	AF317267.1	—	1617160	5918217	—	—	—	—	—
Spathodus marlieri	EF679260	—	1617164	5918215	—	—	—	—	—
Steatocranus bleheri	GQ167789	GQ167978	—	—	GQ168292	GQ168041	GQ168103	GQ168229	GQ168166
Steatocranus casuarius	AF317247.1	GQ167979	—	2394147	GQ168293	GQ168042	GQ168104	GQ168230	GQ168167
Steatocranus gibbiceps	—	GQ167980	—	—	GQ168294	GQ168043	GQ168105	GQ168231	GQ168168
Steatocranus glaber	—	GQ168005	—	—	GQ168319	GQ168067	GQ168130	GQ168256	GQ168193
Steatocranus irvinei	GQ167792	GQ167981	—	—	GQ168295	GQ168044	GQ168106	GQ168232	GQ168169
Steatocranus sp. bulky head	GQ167793	GQ167982	—	—	GQ168296	GQ168045	GQ168107	GQ168233	GQ168170
Steatocranus sp. dwarf	—	GQ167983	—	—	GQ168297	GQ168046	GQ168108	GQ168234	GQ168171

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Steatocranus sp. redeye	GQ167808	GQ167997	—	—	GQ168311	GQ168059	GQ168122	GQ168248	GQ168185
Steatocranus tinanti	AF317248.1	58199027	—	—	GQ168298	GQ168047	GQ168109	GQ168235	GQ168172
Steatocranus ubanguiensis	—	GQ168014	—	—	GQ168329	GQ168077	GQ168140	GQ168266	GQ168203
Stigmatochromis modestus	—	—	62637982	—	—	—	—	—	—
Stigmatochromis woodi	AF305299	—	7595688	—	—	—	—	—	—
Stomatepia mariae	AF317279	18182242	55468940	x	GQ168299	GQ168048	GQ168110	GQ168236	GQ168173
Stomatepia mongo	AJ845094	—	55468938	x	—	—	—	—	—
Stomatepia pindu	AJ845098	18182243	55468942	x	—	—	—	—	—
Taeniochromis holotaenia	—	—	62637987	—	—	—	—	—	—
Taeniolethrinops furcicauda	EF585263	—	62637984	—	—	—	—	—	—
Taeniolethrinops laticeps	AF305305	—	22531787	—	—	—	—	—	—
Taeniolethrinops praeorbitalis	AF305318	—	22531800	—	—	—	—	—	—
Tanganicodus irsacae	AF398219	—	1617161	313093	—	—	—	—	—
Telmatochromis bifrenatus	DQ055009	—	510217	510215	—	—	—	—	—
Telmatochromis brichardi	EF462236	—	—	—	—	—	—	—	—
Telmatochromis burgeoni	—	—	510218	510216	—	—	—	—	—
Telmatochromis dhonti	EF679266	—	47116792	EF679298	—	—	—	—	—
Telmatochromis sp.	—	18182240	—	—	—	—	—	—	—
Telmatochromis temporalis	EF462234	—	47116796	EF679293	—	—	—	—	—
Telmatochromis vittatus	AY740396	—	510222	510221	—	—	—	—	—
Telotrematocara macrostoma	—	—	52221317	52221274	—	—	—	—	—
Thoracochromis albolabris	EU753929.1	—	58866301	—	—	—	—	—	—
Thoracochromis aviumLA	x	—	—	—	—	—	—	—	—
Thoracochromis brauschi	AY930080	GQ168007	30143258	—	GQ168321	GQ168069	GQ168132	GQ168258	GQ168195
Thoracochromis buysi	EU753933.1	—	58866305	—	—	—	—	—	—
Thoracochromis demeusii	—	—	58866311	—	—	—	—	—	—
Thoracochromis mahagiensisLA	x	—	—	—	—	—	—	—	—
Thoracochromis moeruensis black	—	—	x	—	—	—	—	—	—
Thoracochromis moeruensis yellow	—	—	x	—	—	—	—	—	—
Thoracochromis oligacanthus	AF416779.1	—	—	—	—	—	—	—	—
Thoracochromis petroniusLE	x	—	—	—	—	—	—	—	—
Thoracochromis pharyngalisLE	x	—	x	—	—	—	—	—	—
Thoracochromis pundamilia-like	—	—	x	—	—	—	—	—	—
Thoracochromis red spotted fin	—	—	x	—	—	—	—	—	—
Thoracochromis rudolfianus	EU753942.1	—	—	—	—	—	—	—	—
Thoracochromis wingatiiLA	x	—	—	—	—	—	—	—	—
Thysochromis ansorgii	AY663713.1	—	—	—	—	—	—	—	—
Tilapia ap. aff. rheophila Samou	—	GQ168013	—	—	GQ168328	GQ168076	GQ168139	GQ168265	GQ168202
Tilapia bakossi	—	—	x	x	—	—	—	—	—
Tilapia bemini	—	—	15428684	x	—	—	—	—	—
Tilapia bilineata Lefini	GQ167775	GQ167964	—	—	GQ168278	GQ168027	GQ168089	GQ168215	GQ168152
Tilapia bilineata Salonga	—	GQ168012	—	—	GQ168327	GQ168075	GQ168138	GQ168264	GQ168201
Tilapia brevimanus	AF317249.1	GQ168016	—	—	GQ168331	GQ168079	GQ168142	GQ168268	GQ168205
Tilapia busumana	AF317250.1	GQ167987	—	x	GQ168301	GQ168049	GQ168112	GQ168238	GQ168175
Tilapia buttikoferi	AF317251.1	GQ167986	—	—	GQ168300	—	GQ168111	GQ168237	GQ168174
Tilapia cabrae	AF317252.1	—	—	—	—	—	—	—	—
Tilapia cessiana	AF317253.1	—	—	—	—	—	—	—	—
Tilapia cf. nyongana Dja	—	GQ168015	—	—	GQ168330	GQ168078	GQ168141	GQ168267	GQ168204
Tilapia cf. rheophila	GQ167825	—	—	—	—	—	—	—	—
Tilapia coffea	AF317254.1	—	—	—	—	—	—	—	—
Tilapia dageti	GQ167821	GQ168010	—	—	GQ168324	GQ168072	GQ168135	GQ168261	GQ168198
Tilapia deckerti Eja	—	—	x	x	—	—	—	—	—
Tilapia discolor	AF317255.1	GQ167990	15428686	—	GQ168304	GQ168052	GQ168115	GQ168241	GQ168178
Tilapia Eja Jewel	—	—	x	x	—	—	—	—	—
Tilapia Eja large	—	—	x	x	—	—	—	—	—
Tilapia Eja littleone	—	—	x	x	—	—	—	—	—
Tilapia flava	—	—	x	x	—	—	—	—	—
Tilapia guineensis Cross	—	—	x	x	—	—	—	—	—
Tilapia guineensis Ivoire	—	—	x	x	—	—	—	—	—
Tilapia guineensis Nguti	—	—	—	x	—	—	—	—	—
Tilapia guinasana	GQ167802	GQ167991	—	—	GQ168305	GQ168053	GQ168116	GQ168242	GQ168179
Tilapia guineensis	AF317256.1	GQ168025	15428685	—	GQ168340	GQ168088	GQ168151	GQ168277	GQ168214
Tilapia gutturosa	—	—	x	x	—	—	—	—	—
Tilapia imbriferna	—	—	x	x	—	—	—	—	—
Tilapia joka	GQ167803	GQ167992	—	—	GQ168306	GQ168054	GQ168117	GQ168243	GQ168180
Tilapia kottae	—	—	x	x	—	—	—	—	—
Tilapia louka	AF317257.1	GQ168011	—	—	GQ168325	GQ168073	GQ168136	GQ168262	GQ168199
Tilapia mariae	AF317258.1	GQ168026	15428683	x	GQ168326	GQ168074	GQ168137	GQ168263	GQ168200
Tilapia rendalli	AF317259.1	—	15428689	2394117	—	—	—	—	—
Tilapia ruweti	GQ167799	GQ167988	15428692	—	GQ168302	GQ168050	GQ168113	GQ168239	GQ168176
Tilapia snyderae	—	—	x	x	—	—	—	—	—
Tilapia sp. aff. zillii Kisangani	—	GQ168017	—	—	GQ168332	GQ168080	GQ168143	GQ168269	GQ168206
Tilapia sparrmanii	AF317260.1	134268651	15428693	134303319	GQ168303	GQ168051	GQ168114	GQ168240	EF470877
Tilapia tholloni	GQ167804	GQ167993	—	—	GQ168307	GQ168055	GQ168118	GQ168244	GQ168181
Tilapia walteri	AF317261.1	—	—	—	—	—	—	—	—
Tilapia zillii	AF317262.1	GQ168024	15428690	x	GQ168339	GQ168087	GQ168150	GQ168276	GQ168213
Tramitichromis brevis	AF305320	—	22531802	—	—	—	—	—	—
Tramitichromis intermedius	—	—	62637992	—	—	—	—	—	—
Tramitichromis lituris	—	—	13235348	—	—	—	—	—	—
Tramitichromis variabilis	AF305319	—	22531801	—	—	—	—	—	—
Trematocara macrostoma	AY663715.1	—	—	—	—	—	—	—	—
Trematocranus placodon	EF585261	—	62637990	—	—	—	—	—	—
Trematocara unimaculatum	AF317268.1	—	52221316	18265840	—	—	—	—	—
Triglachromis otostigma	AF398217	—	510220	510219	—	—	—	—	—
Tristramella simonis	AF317276.1	GQ168002	—	—	GQ168316	GQ168064	GQ168127	GQ168253	GQ168190
Tropheus annectens	—	—	13736	64099	—	—	—	—	—
Tropheus brichardi	AY930086.1	—	13746	13742	—	—	—	—	—
Tropheus duboisi	AY930085.1	134268655	13757	13751	—	—	—	—	EF470878
Tropheus kasabae	—	—	13779	13775	—	—	—	—	—
Tropheus mooriiLT	x	7576475	1495721	13781	—	—	—	—	—

Species	ND2	16S	CR	CytB	ENCI	Ptr	S7 Intron	SH3PX2	Tmo-4c4
Tropheus polli	AY930084.1	---	13865	13852	---	---	---	---	---
Tylochromis bangwelensis	---	---	112735228	---	---	---	---	---	---
Tylochromis cf. variabilis MK-2006	---	---	112735226	---	---	---	---	---	---
Tylochromis lateralis	---	---	112735227	---	---	---	---	---	---
Tylochromis leonensis	AF317274.1	33090449	---	---	---	---	---	---	---
Tylochromis mylodon	---	---	112735229	---	---	---	---	---	---
Tylochromis polylepis	AB018973.2	5114133	112735238	18265850	---	---	---	---	U70337
Tylochromis pulcher	---	58199028	---	---	---	---	---	---	---
Tylochromis sp.	---	---	---	---	GQ168312	GQ168060	GQ168123	GQ168249	GQ168186
Tyrannochromis nigriventer	AF305307	---	22531789	---	---	---	---	---	---
Variabilichromis moorii	DQ055016	18182236	85681972	510172	GQ168313	GQ168061	GQ168124	GQ168250	GQ168187
Xenochromis hecqui	EF437514.1	---	---	---	---	---	---	---	---
Xenotilapia bathyphila	AY337789.1	---	33355557	33356090	---	---	---	---	---
Xenotilapia boulengeri	---	---	33355559	33356050	---	---	---	---	---
Xenotilapia caudafasciata	AY337777.1	---	33355565	33356034	---	---	---	---	---
Xenotilapia flavipinnis	AY337794.1	---	33355560	33356026	---	---	---	---	---
Xenotilapia longispinis	AY337778.1	---	33355567	33356038	---	---	---	---	---
Xenotilapia melanogenys	AY682517.1	---	---	---	---	---	---	---	---
Xenotilapia ochrogenys	AY337767.1	---	313101	313099	---	---	---	---	---
Xenotilapia ornatipinnis	---	18182241	---	---	---	---	---	---	---
Xenotilapia papilio	AY337776.1	---	---	---	---	---	---	---	---
Xenotilapia sima	AY337785.1	---	18029953	33356008	---	---	---	---	---
Xenotilapia sp. papilio sunflower	---	---	33355574	33356022	---	---	---	---	---
Xenotilapia spiloptera	AY337788.1	---	33355570	33356032	---	---	---	---	---
Xystichromis phytophagus	AY930076.1	---	51320047	---	---	---	---	---	---
Yssichromis laparogrammaLV	x	---	7595582	---	---	---	---	---	---
Yssichromis pyrrhocephalusLV	x	18182246	AB439318	---	---	---	---	---	---
500 A870MPC18	x	---	---	---	---	---	---	---	---
503 A994MPK8	x	---	---	---	---	---	---	---	---

Supplementary Table 2. Lake physical and environmental variables used in this study.

Lake	Latitude	Longitude	Lake SA (km2)	Lake depth (m)	Surface Elevation (m)	Energy (W/m2/day)	Predator presence/ absence
Abaya	6.43	37.88	1162.0	13.0	1285	49.79	1
Abbe	11.17	41.75	250.0	37.0	243	53.46	0
Albert	1.67	30.92	5300.0	58.0	615	79.92	1
Bangweulu	-11.08	29.75	1510.0	10.0	1140	65.59	1
Baringo	0.63	36.08	129.0	2.1	970	63.57	0
Barombi Mbo	4.67	9.40	4.2	111.0	315	60.26	0
Barombi ba Koto	4.47	9.26	1.2	6.2	100	61.18	0
Bermin	5.16	9.64	0.6	--	530	67.24	0
Bosumtwi	6.51	-1.41	52.0	76.0	150	59.90	--
Chad	13.00	14.00	9400.0	10.5	280	31.78	1
Chala	-3.32	37.70	4.2	94.0	880	34.91	0
Chamo	5.83	37.55	450.0	13.0	1235	50.52	0
Chilwa	-15.30	35.70	1750.0	2.7	622	34.91	0
Chiuta	-14.85	35.85	35.0	3.5	631	37.30	0
Debo	15.32	-4.10	160.0	3.0	259	27.01	--
Edward	-0.33	29.60	2325.0	112.0	912	72.39	0
Ejagham	5.74	8.99	0.5	18.0	129	69.82	0
Eyasi	-3.63	35.08	1160.0	1.0	1037	51.99	0
Fwa	-5.73	23.34	1.7	22.0	680	73.67	0
Guinas sink hole	-19.23	17.36	0.0	130.0	1224	40.79	0
Jipe	-3.61	37.76	30.0	3.0	706	33.99	0
Kagera lakes	-1.62	30.95	86.4	--	1382	55.67	0
Kinneret	32.83	35.58	168.0	43.0	-209	2.20	0
Kivu	-2.00	29.00	2370.0	485.0	1460	64.67	0
Lake Mai Ndombe	-2.00	18.30	2300.0	10.0	291	80.66	1
Lake Tschungruru/Masoko	-9.33	33.76	0.4	--	860	47.22	0
Langano	7.60	38.72	240.0	46.0	1584	51.63	0
Lutoto	-0.33	30.10	0.7	--	1409	65.59	0
Malawi	-12.18	34.37	22490.0	706.0	500	54.57	0
Manyara	-3.58	35.83	250.0	3.7	953	45.38	0
Mareotis	31.15	29.90	110.0	--	-4	10.29	1
Mweru	-9.00	28.75	5040.0	37.0	922	69.63	1
Nabugabo	-0.37	31.89	25.0	5.0	1138	65.04	0
Natron	-2.42	36.00	750.0	3.0	602	51.99	0
Nshere	-0.19	30.14	3.0	--	1022	66.14	0
Rukwa	-8.00	32.42	5700.0	1.0	793	62.65	1
Saka	0.69	30.24	1.4	12.0	1565	69.82	0
Stefani	4.72	36.95	915.0	3.0	1880	50.71	0
Tana	12.00	37.25	3156.0	14.0	1788	64.49	0
Tanganyika	-6.50	29.50	32900.0	1470.0	773	67.06	1
Tumba	-0.83	18.00	750.0	6.0	294	81.21	1
Turkana	3.05	36.02	7300.0	100.0	360.4	60.26	1
Upemba lakes	-8.60	26.43	530.0	4.0	580	67.43	1
Victoria	-1.00	33.00	68800.0	69.0	1133	86.17	0
Zwai	8.00	38.83	430.0	8.0	1846	50.34	0
Logipi	2.23	36.57	72.0	5.0	312	59.90	--

Supplementary Table 3. Diversification states and trait data for cichlid lineages included in this study.

Colonizing taxon	Lake	Diversifying (threshold 1)	Mating system	Mouthbrooding	Egg spots	Haplochromine egg spots	Strongly sexually dimorphic
Oreochromis niloticus	Abaya	no	polygamous	mouthbrooding	no	no	no
Oreochromis niloticus	Abbe	no	polygamous	mouthbrooding	no	no	no
Hemichromis fasciatus	Albert	no	not polygamous	non-mouthbrooding	no	no	no
Oreochromis niloticus-leucostictus	Albert	yes	polygamous	mouthbrooding	no	no	no
Pseudocrenilabrus multicolor	Albert	no	polygamous	mouthbrooding	yes	no	yes
Sarotherodon galilaeus	Albert	no	polygamous	mouthbrooding	no	no	no
Thoracochromis	Albert	yes	polygamous	mouthbrooding	yes	yes	yes
Tilapia zilii	Albert	no	not polygamous	non-mouthbrooding	no	no	no
Hemichromis	Bangweulu	no	not polygamous	non-mouthbrooding	no	no	no
Oreochromis macrochir	Bangweulu	no	polygamous	mouthbrooding	no	no	yes
Pseudocrenilabrus philander	Bangweulu	no	polygamous	mouthbrooding	yes	no	yes
Sargochromis mellandi	Bangweulu	no	polygamous	mouthbrooding	yes	yes	yes
Serranochromis altus	Bangweulu	no	polygamous	mouthbrooding	yes	yes	yes
Serranochromis angusticeps	Bangweulu	no	polygamous	mouthbrooding	yes	yes	yes
Serranochromis robustus	Bangweulu	no	polygamous	mouthbrooding	yes	yes	yes
Serranochromis thumbergi	Bangweulu	no	polygamous	mouthbrooding	yes	yes	yes
Tilapia rendalli	Bangweulu	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia sparmanii	Bangweulu	no	not polygamous	non-mouthbrooding	no	no	0
Tylochromis bangwelensis	Bangweulu	no	polygamous	mouthbrooding	no	no	1
Oreochromis niloticus	Baringo	no	polygamous	mouthbrooding	no	no	0
Chromidotilapia	Barombi ba Koto	no	not polygamous	mouthbrooding	no	no	0
Sarotherodon galilaeus	Barombi ba Koto	no	polygamous	mouthbrooding	no	no	0
Tilapia guineensis	Barombi ba Koto	no	not polygamous	non-mouthbrooding	no	no	0
Sarotherodon galilaeus	Barombi Mbo	yes	polygamous	mouthbrooding	no	no	0
Tilapia guineensis complex	Bermin	yes	not polygamous	non-mouthbrooding	no	no	0
Sarotherodon galilaeus multifasciatus	Bosumtwi	no	polygamous	mouthbrooding	no	no	0
Hemichromis frempongi	Bosumtwi	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia busumana	Bosumtwi	no	not polygamous	non-mouthbrooding	no	no	0
Astatotilapia sp.	Chad	no	polygamous	mouthbrooding	yes	yes	1
Hemichromis fasciatus	Chad	no	not polygamous	non-mouthbrooding	no	no	0
Hemichromis letournauxi	Chad	no	not polygamous	non-mouthbrooding	no	no	0
Oreochromis aureus	Chad	no	polygamous	mouthbrooding	no	no	0
Oreochromis niloticus	Chad	no	polygamous	mouthbrooding	no	no	0
Sarotherodon galilaeus	Chad	no	polygamous	mouthbrooding	no	no	0
Tilapia dageti	Chad	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia zilii	Chad	no	not polygamous	non-mouthbrooding	no	no	0
Astatotilapia bloyeti	Chala	no	polygamous	mouthbrooding	yes	yes	0
Oreochromis pangani	Chala	yes	polygamous	mouthbrooding	no	no	0
Oreochromis niloticus	Chamo	no	polygamous	mouthbrooding	no	no	0
Astatotilapia calliptera	Chilwa	no	polygamous	mouthbrooding	yes	yes	1
Astatotilapia tweddlei	Chilwa	no	polygamous	mouthbrooding	yes	yes	0
Oreochromis shiranus chilwae	Chilwa	no	polygamous	mouthbrooding	no	no	1
Pseudocrenilabrus philander	Chilwa	no	polygamous	mouthbrooding	yes	no	1
Tilapia rendalli	Chilwa	no	not polygamous	non-mouthbrooding	no	no	0
Astatotilapia calliptera	Chiuta	no	polygamous	mouthbrooding	yes	yes	1
Astatotilapia tweddlei	Chiuta	no	polygamous	mouthbrooding	yes	yes	0
Oreochromis shiranus chilwae	Chiuta	no	polygamous	mouthbrooding	no	no	1
Pseudocrenilabrus philander	Chiuta	no	polygamous	mouthbrooding	yes	no	1
Tilapia rendalli	Chiuta	no	not polygamous	non-mouthbrooding	no	no	0
Astatotilapia sp.	Debo	no	polygamous	mouthbrooding	yes	yes	1
Hemichromis fasciatus	Debo	no	not polygamous	non-mouthbrooding	no	no	0
Oreochromis niloticus	Debo	no	polygamous	mouthbrooding	no	no	0
Sarotherodon galilaeus	Debo	no	polygamous	mouthbrooding	no	no	0
Astatoreochromis alluaudi	Edward	no	polygamous	mouthbrooding	yes	yes	0
Haplochromis	Edward	yes	polygamous	mouthbrooding	yes	yes	1
Pseudocrenilabrus multicolor victoriae	Edward	no	polygamous	mouthbrooding	yes	no	1
Thoracochromis	Edward	yes	polygamous	mouthbrooding	yes	yes	1
Sarotherodon galilaeus	Ejagham	yes	polygamous	mouthbrooding	no	no	0
Tilapia guineensis complex	Ejagham	yes	not polygamous	non-mouthbrooding	no	no	0
Astatotilapia cf. *nubila*	Eyasi	no	polygamous	mouthbrooding	yes	yes	1
Oreochromis amphimelas	Eyasi	no	polygamous	mouthbrooding	no	no	1
Thoracochromis	Fwa	yes	polygamous	mouthbrooding	no_yes	no	1
Tilapia guinasana	Guinas sink hole	yes	not polygamous	non-mouthbrooding	no	no	0
Oreochromis pangani	Jipe	yes	polygamous	mouthbrooding	no	no	1
Astatoreochromis alluaudi	Kagera lakes	no	polygamous	mouthbrooding	yes	yes	0
Haplochromis	Kagera lakes	yes	polygamous	mouthbrooding	yes	yes	1
Oreochromis esculentus	Kagera lakes	no	polygamous	mouthbrooding	no	no	1
Astatotilapia flavijosephi	Kinneret	no	polygamous	mouthbrooding	yes	yes	1
Oreochromis aureus	Kinneret	no	polygamous	mouthbrooding	no	no	0
Oreochromis niloticus	Kinneret	no	polygamous	mouthbrooding	no	no	0
Sarotherodon galilaeus	Kinneret	no	polygamous	mouthbrooding	no	no	0
Tilapia zilii	Kinneret	no	not polygamous	non-mouthbrooding	no	no	0
Tristramella	Kinneret	yes	--	mouthbrooding	no	no	--
Astatotilapia gracilior	Kivu	no	polygamous	mouthbrooding	yes	yes	1
Haplochromis	Kivu	yes	polygamous	mouthbrooding	yes	yes	1
Oreochromis niloticus	Kivu	no	polygamous	mouthbrooding	no	no	0
Hemichromis cerasogaster	Lake Mai Ndombe	no	not polygamous	non-mouthbrooding	no	no	0
Hemichromis fasciatus	Lake Mai Ndombe	no	not polygamous	non-mouthbrooding	no	no	0
Nanochromis	Lake Mai Ndombe	yes	not polygamous	non-mouthbrooding	no	no	0
Nanochromis sp. *"green speckle"*	Lake Mai Ndombe	no	not polygamous	non-mouthbrooding	no	no	0
Pelmatochromis nigrofasciatus	Lake Mai Ndombe	no	not polygamous	non-mouthbrooding	no	no	0
Tylochromis	Lake Mai Ndombe	no	polygamous	mouthbrooding	no	no	--
Oreochromis chungruruensis	Lake Tschungruru/Masok	yes	polygamous	mouthbrooding	no	no	1
Oreochromis niloticus	Langano	no	polygamous	mouthbrooding	no	no	1
Oreochromis niloticus	Logipi	no	polygamous	mouthbrooding	no	no	1
Haplochromis sp. *Lutoto*	Lutoto	no	polygamous	mouthbrooding	yes	yes	1
Astatotilapia	Malawi	yes	polygamous	mouthbrooding	yes	yes	1
Oreochromis (Nyasalapia)	Malawi	yes	polygamous	mouthbrooding	no	no	1
Oreochromis shiranus	Malawi	no	polygamous	mouthbrooding	no	no	1

Pseudocrenilabrus philander	Malawi	no	polygamous	mouthbrooding	yes	no	1
Serranochromis robustus	Malawi	no	polygamous	mouthbrooding	yes	yes	1
Tilapia rendalli	Malawi	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia sparrmanii	Malawi	no	not polygamous	non-mouthbrooding	no	no	0
Astatotilapia sparsidens	Manyara	no	polygamous	mouthbrooding	yes	yes	1
Oreochromis amphimelas	Manyara	no	polygamous	mouthbrooding	no	no	1
Astatotilapia sp.	Mareotis	no	polygamous	mouthbrooding	yes	yes	1
Hemichromis letournauxi	Mareotis	no	not polygamous	non-mouthbrooding	no	no	0
Pseudocrenilabrus multicolor	Mareotis	no	polygamous	mouthbrooding	yes	no	1
Hemichromis elongatus	Mweru	no	not polygamous	non-mouthbrooding	no	no	0
Oroechromis macrochir	Mweru	no	polygamous	mouthbrooding	no	no	1
Orthochromis kalungwishiensis	Mweru	no	--	mouthbrooding	no	no	0
Orthochromis polyacanthus	Mweru	no	--	mouthbrooding	no	no	0
Pseudocrenilabrus	Mweru	yes	polygamous	mouthbrooding	yes	no	1
Sargochromis	Mweru	yes	polygamous	mouthbrooding	yes	yes	1
Serranochromis "large tooth"	Mweru	yes	polygamous	mouthbrooding	yes	yes	1
Serranochromis "small tooth"	Mweru	yes	polygamous	mouthbrooding	yes	yes	1
Serranochromis robustus	Mweru	no	polygamous	mouthbrooding	yes	yes	1
Tilapia rendalli	Mweru	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia sparmanii	Mweru	no	not polygamous	non-mouthbrooding	no	no	0
Tylochromis mylodon	Mweru	no	polygamous	mouthbrooding	no	no	1
Astatoreochromis alluaudi	Nabugabo	no	polygamous	mouthbrooding	yes	yes	0
Haplochromis	Nabugabo	yes	polygamous	mouthbrooding	yes	yes	1
Oreochromis variabilis	Nabugabo	no	polygamous	mouthbrooding	no	no	1
Oreochromis esculentus	Nabugabo	no	polygamous	mouthbrooding	no	no	1
Pseudocrenilabrus multicolor victoriae	Nabugabo	no	polygamous	mouthbrooding	yes	no	1
Oreochromis alcalicus	Natron	yes	polygamous	mouthbrooding	no	no	1
Haplochromis sp. *Nshere*	Nshere	no	polygamous	mouthbrooding	yes	yes	1
Astatotilapia pseudopaludinosa	Rukwa	no	polygamous	mouthbrooding	yes	yes	1
Haplochromis	Rukwa	yes	polygamous	mouthbrooding	yes	yes	1
Oreochromis rukwaensis	Rukwa	no	polygamous	mouthbrooding	no	no	1
Pseudocrenilabrus sp. aff. *philander*	Rukwa	no	polygamous	mouthbrooding	yes	no	1
Tilapia rendalli	Rukwa	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia sp. aff. *sparrmanii*	Rukwa	no	not polygamous	non-mouthbrooding	no	no	0
Astatoreochromis alluaudi	Saka	no	polygamous	mouthbrooding	yes	yes	0
Haplochromis	Saka	yes	polygamous	mouthbrooding	yes	yes	1
Pseudocrenilabrus multicolor victoriae	Saka	no	polygamous	mouthbrooding	yes	no	1
Oreochromis niloticus	Stefani	no	polygamous	mouthbrooding	no	no	0
Oreochromis niloticus	Tana	no	polygamous	mouthbrooding	no	no	0
Astatoreochromis	Tanganyika	no	polygamous	mouthbrooding	yes	yes	0
Bathybates	Tanganyika	yes	--	mouthbrooding	yes	no	0
Boulengerochromis	Tanganyika	no	not polygamous	non-mouthbrooding	no	no	0
Cyprichromini	Tanganyika	yes	polygamous	mouthbrooding	no	no	1
Ectodini	Tanganyika	yes	polygamous	mouthbrooding	no	no	1
Eretmodini	Tanganyika	yes	not polygamous	mouthbrooding	no	no	0
Lamprologini	Tanganyika	yes	not polygamous	non-mouthbrooding	no	no	0
Limnochromis	Tanganyika	yes	not polygamous	mouthbrooding	no	no	0
Oreochromis niloticus	Tanganyika	no	polygamous	mouthbrooding	no	no	0
Oreochromis tanganicae	Tanganyika	no	polygamous	mouthbrooding	no	no	0
Perissodini	Tanganyika	yes	not polygamous	mouthbrooding	no	no	0
Pseudocrenilabrus sp.	Tanganyika	no	polygamous	mouthbrooding	yes	no	1
Tilapia rendalli	Tanganyika	no	not polygamous	non-mouthbrooding	no	no	0
Trematocara	Tanganyika	yes	--	mouthbrooding	no	no	0
Tropheini	Tanganyika	yes	polygamous	mouthbrooding	yes	no	0
Tylochromis polylepis	Tanganyika	no	polygamous	mouthbrooding	no	no	--
Hemichromis elongatus	Tumba	no	not polygamous	non-mouthbrooding	no	no	0
Lamprologus tumbanus	Tumba	no	not polygamous	non-mouthbrooding	no	no	0
Nanochromis	Tumba	no	not polygamous	non-mouthbrooding	no	no	0
Tylochromis	Tumba	no	polygamous	mouthbrooding	no	no	--
Tylochromis microdon	Tumba	no	polygamous	mouthbrooding	no	no	--
Hemichromis letournauxi	Turkana	no	not polygamous	non-mouthbrooding	no	no	0
Oreochromis niloticus	Turkana	no	polygamous	mouthbrooding	no	no	0
Sarotherodon galilaeus	Turkana	no	polygamous	mouthbrooding	no	no	0
Thoracochromis	Turkana	yes	polygamous	mouthbrooding	yes	yes	1
Tilapia zilii	Turkana	no	not polygamous	non-mouthbrooding	no	no	0
Lamprologus symoensis	Upemba lakes	no	not polygamous	non-mouthbrooding	no	no	0
Oreochromis upembae	Upemba lakes	no	polygamous	mouthbrooding	no	no	1
Pseudocrenilabrus nicholsi	Upemba lakes	no	polygamous	mouthbrooding	yes	no	1
Tilapia rendalli	Upemba lakes	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia ruweti	Upemba lakes	no	not polygamous	non-mouthbrooding	no	no	0
Tilapia sparmanii	Upemba lakes	no	not polygamous	non-mouthbrooding	no	no	0
Astatoreochromis alluaudi	Victoria	no	polygamous	mouthbrooding	yes	yes	0
Haplochromis	Victoria	yes	polygamous	mouthbrooding	yes	yes	1
Oreochromis esculentus	Victoria	no	polygamous	mouthbrooding	no	no	1
Oreochromis variabilis	Victoria	no	polygamous	mouthbrooding	no	no	1
Pseudocrenilabrus multicolor victoriae	Victoria	no	polygamous	mouthbrooding	yes	no	1
Oreochromis niloticus	Zwai	no	polygamous	mouthbrooding	no	no	0

Supplementary Table 4. Times for divergence, calculated either as midpoint of the age range estimate for the most recent lake desiccation, mean stem age of the clade calculated from calibrated molecular phylogenies (see Methods, SI section 1 for details), or calibrated molecular dates from other studies.

Radiating taxon	Lake	Time for Divergence	References
Oreochromis niloticus	Abaya	--	
Oreochromis niloticus	Abbe	0.019	Barker & Gasse 2003
Thoracochromis	Albert	0.015	Beuning et al. 1997
Hemichromis fasciatus	Albert	0.015	Beuning et al. 1997
Oreochromis niloticus-leucostictus	Albert	0.015	Beuning et al. 1997
Pseudocrenilabrus multicolor	Albert	0.015	Beuning et al. 1997
Sarotherodon galilaeus	Albert	0.015	Beuning et al. 1997
Tilapia zilii	Albert	0.015	Beuning et al. 1997
Hemichromis	Bangweulu	0.008	Scholz et al. 2007
Oreochromis macrochir	Bangweulu	0.008	Scholz et al. 2007
Pseudocrenilabrus philander	Bangweulu	0.008	Scholz et al. 2007
Sargochromis mellandi	Bangweulu	0.008	Scholz et al. 2007
Serranochromis altus	Bangweulu	0.008	Scholz et al. 2007
Serranochromis angusticeps	Bangweulu	0.008	Scholz et al. 2007
Serranochromis robustus	Bangweulu	0.008	Scholz et al. 2007
Serranochromis thumbergi	Bangweulu	0.008	Scholz et al. 2007
Tilapia rendalli	Bangweulu	0.008	Scholz et al. 2007
Tilapia sparmanii	Bangweulu	0.008	Scholz et al. 2007
Tylochromis bangwelensis	Bangweulu	0.008	Scholz et al. 2007
Oreochromis niloticus	Baringo	0.015	Owen et al. 1982; Garcin et al. 2009; Bergner et al. 2009
Chromidotilapia	Barombi ba Koto	--	
Sarotherodon galilaeus	Barombi ba Koto	--	
Tilapia guineensis	Barombi ba Koto	--	
Sarotherodon galilaeus	Barombi Mbo	0.900 **	
Tilapia guineensis complex	Bermin	0.750 *	Schliewen et al. 1994
Hemichromis frempongi	Bosumtwi	0.070	Scholz et al. 2007
Sarotherodon galilaeus multifasciatus	Bosumtwi	0.070	Scholz et al. 2007
Tilapia busumana	Bosumtwi	0.070	Scholz et al. 2007
Astatotilapia sp.	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Hemichromis fasciatus	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Hemichromis letournauxi	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Oreochromis aureus	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Oreochromis niloticus	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Sarotherodon galilaeus	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Tilapia dageti	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Tilapia zilii	Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Astatotilapia bloyeti	Chala	0.140	Moernaut et al. 2010
Oreochromis pangani	Chala	0.140	Moernaut et al. 2010
Oreochromis niloticus	Chamo	--	
Astatotilapia calliptera	Chilwa	0.008	Scholz et al. 2007
Astatotilapia tweddlei	Chilwa	0.008	Scholz et al. 2007
Oreochromis shiranus chilwae	Chilwa	0.008	Scholz et al. 2007
Pseudocrenilabrus philander	Chilwa	0.008	Scholz et al. 2007
Tilapia rendalli	Chilwa	0.008	Scholz et al. 2007
Astatotilapia calliptera	Chiuta	0.008	Scholz et al. 2007
Astatotilapia tweddlei	Chiuta	0.008	Scholz et al. 2007
Oreochromis shiranus chilwae	Chiuta	0.008	Scholz et al. 2007
Pseudocrenilabrus philander	Chiuta	0.008	Scholz et al. 2007
Tilapia rendalli	Chiuta	0.008	Scholz et al. 2007
Astatotilapia sp.	Debo	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Hemichromis fasciatus	Debo	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Oreochromis niloticus	Debo	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Sarotherodon galilaeus	Debo	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Astatoreochromis alluaudi	Edward	0.270	Beuning et al. 1997; Laerdal & Talbot 2002
Haplochromis	Edward	0.270	Beuning et al. 1997; Laerdal & Talbot 2002
Pseudocrenilabrus multicolor victoriae	Edward	0.270	Beuning et al. 1997; Laerdal & Talbot 2002
Thoracochromis	Edward	0.270	Beuning et al. 1997; Laerdal & Talbot 2002
Sarotherodon galilaeus	Ejagham	0.010	Schliewen et al. 2001
Tilapia guineensis complex	Ejagham	0.010	Schliewen et al. 2001
Astatotilapia cf. nubila	Eyasi	0.015	Johnson et al. 2000; Scholz et al. 2007 ; Stager & Johnson 2008
Oreochromis amphimelas	Eyasi	0.015	Johnson et al. 2000; Scholz et al. 2007 ; Stager & Johnson 2008
Thoracochromis	Fwa	--	
Tilapia guinasana	Guinas sink hole	--	
Oreochromis pangani	Jipe	--	
Astatoreochromis alluaudi	Kagera lakes	--	
Haplochromis	Kagera lakes	--	
Oreochromis esculentus	Kagera lakes	--	
Astatotilapia flavijosephi	Kinneret	0.070	Hazan et al. 2005
Oreochromis aureus	Kinneret	0.070	Hazan et al. 2005
Oreochromis niloticus	Kinneret	0.070	Hazan et al. 2005
Sarotherodon galilaeus	Kinneret	0.070	Hazan et al. 2005
Tilapia zilii	Kinneret	0.070	Hazan et al. 2005
Tristramella	Kinneret	0.070	Hazan et al. 2005
Haplochromis	Kivu	0.008	Haberyan & Hecky, 1987
Astatotilapia gracilior	Kivu	0.008	Haberyan & Hecky, 1987
Oreochromis niloticus	Kivu	0.008	Haberyan & Hecky, 1987

Radiating taxon	Lake	Time for Divergence	References
Hemichromis cerasogaster	Lake Mai Ndombe	--	
Hemichromis fasciatus?	Lake Mai Ndombe	--	
Nanochromis	Lake Mai Ndombe	--	
Nanochromis sp. "green speckle"	Lake Mai Ndombe	--	
Pelmatochromis nigrofasciatus	Lake Mai Ndombe	--	
Tylochromis	Lake Mai Ndombe	--	
Oreochromis chungruruensis	Lake Tschungruru/Masoko	0.050	Barker & Gasse 2003
Oreochromis niloticus	Langano	0.030	Street 1979; Barker & Gasse 2003
Oreochromis niloticus	Logipi	--	
Haplochromis sp. Lutoto	Lutoto	0.050	Sato et al. 2003
Astatotilapia	Malawi	9.765 **	
Oreochromis shiranus	Malawi	--	
Oreochromis (Nyasalapia)	Malawi	--	
Pseudocrenilabrus philander	Malawi	--	
Serranochromis robustus	Malawi	0.630 *	Genner et al. 2007
Tilapia rendalli	Malawi	--	
Tilapia sparrmanii	Malawi	--	
Astatotilapia sparsidens	Manyara	0.015	Johnson et al. 2000; Scholz et al. 2007 ; Stager & Johnson 2008
Oreochromis amphimelas	Manyara	0.015	Johnson et al. 2000; Scholz et al. 2007 ; Stager & Johnson 2008
Astatotilapia sp.	Mareotis	--	
Hemichromis letournauxi	Mareotis	--	
Pseudocrenilabrus multicolor	Mareotis	--	
Hemichromis elongatus	Mweru	0.045	Scholz et al. 2007
Oroechromis macrochir	Mweru	0.045	Scholz et al. 2007
Orthochromis kalungwishiensis	Mweru	0.045	Scholz et al. 2007
Orthochromis polyacanthus	Mweru	0.045	Scholz et al. 2007
Pseudocrenilabrus	Mweru	0.045	Scholz et al. 2007
Sargochromis	Mweru	0.045	Scholz et al. 2007
Serranochromis "large tooth"	Mweru	0.045	Scholz et al. 2007
Serranochromis robustus	Mweru	0.045	Scholz et al. 2007
Serranochromis "small tooth"	Mweru	0.045	Scholz et al. 2007
Tilapia rendalli	Mweru	0.045	Scholz et al. 2007
Tilapia sparmanii	Mweru	0.045	Scholz et al. 2007
Tylochromis mylodon	Mweru	0.045	Scholz et al. 2007
Astatoreochromis alluaudi	Nabugabo	0.006	Stager et al. 2005
Haplochromis	Nabugabo	0.006	Stager et al. 2005
Oreochromis esculentus	Nabugabo	0.006	Stager et al. 2005
Oreochromis variabilis	Nabugabo	0.006	Stager et al. 2005
Pseudocrenilabrus multicolor victoriae	Nabugabo	0.006	Stager et al. 2005
Oreochromis alcalicus	Natron	0.015	Barker & Gasse 2003
Haplochromis sp. nshere	Nshere	0.050	Sato et al. 2003
Haplochromis	Rukwa	0.015	Barker & Gasse 2003
Astatotilapia pseudopaludinosa	Rukwa	0.015	Barker & Gasse 2003
Oreochromis rukwaensis	Rukwa	0.015	Barker & Gasse 2003
Pseudocrenilabrus sp. aff. philander	Rukwa	0.015	Barker & Gasse 2003
Tilapia rendalli	Rukwa	0.015	Barker & Gasse 2003
Tilapia sp. aff. sparrmanii	Rukwa	0.015	Barker & Gasse 2003
Astatoreochromis alluaudi	Saka	0.003	Bessems et al., 2003; Russell et al., 2007
Haplochromis	Saka	0.003	Bessems et al., 2003; Russell et al., 2007
Pseudocrenilabrus multicolor victoriae	Saka	0.003	Bessems et al., 2003; Russell et al., 2007
Oreochromis niloticus	Stefani	--	
Oreochromis niloticus	Tana	0.017	Lamb et al. 2007
Astatoreochromis	Tanganyika	4.500 *	Koblmueller et al. 2008
Bathybates	Tanganyika	26.036 **	
Boulengerochromis	Tanganyika	34.046 **	
Cyprichromini	Tanganyika	15.197 **	
Ectodini	Tanganyika	22.947 **	
Eretmodini	Tanganyika	26.412 **	
Limnochromis	Tanganyika	18.386 **	
Lamprologini	Tanganyika	26.460 **	
Oreochromis niloticus	Tanganyika	--	
Oreochromis tanganicae	Tanganyika	1.375 **	
Perissodini	Tanganyika	15.437 **	
Pseudocrenilabrus sp.	Tanganyika	7.000 *	Koblmüller et al. 2008
Tilapia rendalli	Tanganyika	--	
Trematocara	Tanganyika	26.043 **	
Tropheini	Tanganyika	15.659 **	
Tylochromis polylepis	Tanganyika	0.500 *	Koch et al. 2007
Hemichromis elongatus	Tumba	--	
Nanochromis	Tumba	--	
Lamprologus tumbanus	Tumba	--	
Tylochromis	Tumba	--	
Tylochromis microdon	Tumba	--	
Thoracochromis	Turkana	0.015	Owen et al. 1982; Beadle 1974; Garcin et al. 2009; Bergner et al. 2009
Hemichromis letournauxi	Turkana	0.015	Owen et al. 1982; Beadle 1974; Garcin et al. 2009; Bergner et al. 2009
Oreochromis niloticus	Turkana	0.015	Owen et al. 1982; Beadle 1974; Garcin et al. 2009; Bergner et al. 2009
Sarotherodon galilaeus	Turkana	0.015	Owen et al. 1982; Beadle 1974; Garcin et al. 2009; Bergner et al. 2009
Tilapia zilii	Turkana	0.015	Owen et al. 1982; Beadle 1974; Garcin et al. 2009; Bergner et al. 2009
Lamprologus symoensis	Upemba lakes	--	
Oreochromis upembae	Upemba lakes	--	
Pseudocrenilabrus nicholsi	Upemba lakes	--	

Radiating taxon	Lake	Time for Divergence	References
Tilapia rendalli	Upemba lakes	--	
Tilapia ruweti	Upemba lakes	--	
Tilapia sparmanii	Upemba lakes	--	
Astatoreochromis alluaudi	Victoria	0.015	Johnson et al. 2000
Haplochromis	Victoria	0.015	Johnson et al. 2000
Oreochromis esculentus	Victoria	0.015	Johnson et al. 2000 ; Stager & Johnson 2008
Oreochromis variabilis	Victoria	0.015	Johnson et al. 2000 ; Stager & Johnson 2008
Pseudocrenilabrus multicolor victoriae	Victoria	0.015	Johnson et al. 2000 ; Stager & Johnson 2008
Oreochromis niloticus	Zwai	0.030	Street 1979; Barker & Gasse 2003

*derived from calibrated molecular divergence date from the provided reference

** derived from our calibrated molecular phylogenies, see Methods and SI section 1 for details

Supplementary Table 6. Single predictor variable phylogenetic logistic regression reveals significant associations between cichlid diversification and lake depth, predators, mouthbrooding, sexual dichromatism, time for diversification, and lake age. When Lake Tanganyika is excluded the same variables are significant, and additionally there is are significant associations between diversification and latitude and diversification and egg spots. N = sample size. a = phylogenetic signal in the diversification state/factor regression.

Predictor	Full Dataset						Excluding Tanganyika					
Environmental Variables	N	Estimate	SE (±)	z-value	z-value p	*a*	N	Estimate	SE (±)	z-value	z-value p	*a*
Lake Surface Area	166	-0.062	0.043	-1.436	0.075	-2.365	150	-0.136	0.059	-2.329	**0.010**	-2.491
Lake Depth	156	0.342	0.090	3.794	**0.000**	-2.615	140	0.317	0.120	2.641	**0.004**	-2.956
Energy	166	0.010	0.009	1.080	0.140	-3.324	150	0.008	0.010	0.830	0.203	-3.520
Latitude	166	-0.043	0.027	-1.610	0.054	-3.396	150	-0.052	0.031	-1.670	**0.047**	-4.000
Latitude (residual)	166	-0.021	0.037	-0.575	0.283	-3.286	150	-0.032	0.041	-0.774	0.220	-3.362
Elevation	166	0.000	0.000	0.767	0.222	-3.457	150	0.000	0.000	0.640	0.261	-3.613
Predators	158	-0.532	0.311	-1.708	**0.044**	-2.719	142	-0.921	0.400	-2.304	**0.011**	-2.860
Traits												
Polygamous Mating System	161	0.472	0.494	0.957	0.169	-3.200	148	0.973	0.611	1.592	0.056	-3.178
Mouthbrooding	166	1.158	0.554	2.091	0.018	-2.905	151	0.937	0.605	1.549	0.061	-3.207
Egg spots	166	0.349	0.443	0.786	0.216	-3.368	151	0.858	0.457	1.877	**0.030**	-4.000
Haplochromine egg spots	166	0.327	0.452	0.722	0.235	-3.456	151	1.071	0.461	2.325	**0.010**	-4.000
Sexual Dichromatism	161	0.744	0.427	1.744	**0.041**	-2.656	147	1.284	0.520	2.468	**0.007**	-2.937
Time												
Time in Lake	126	0.388	0.096	4.034	**0.000**	-2.838	113	0.461	0.176	2.619	**0.004**	-2.668
Lake Age	132	0.246	0.073	3.391	**0.000**	-2.784	117	0.241	0.108	2.232	**0.013**	-2.778

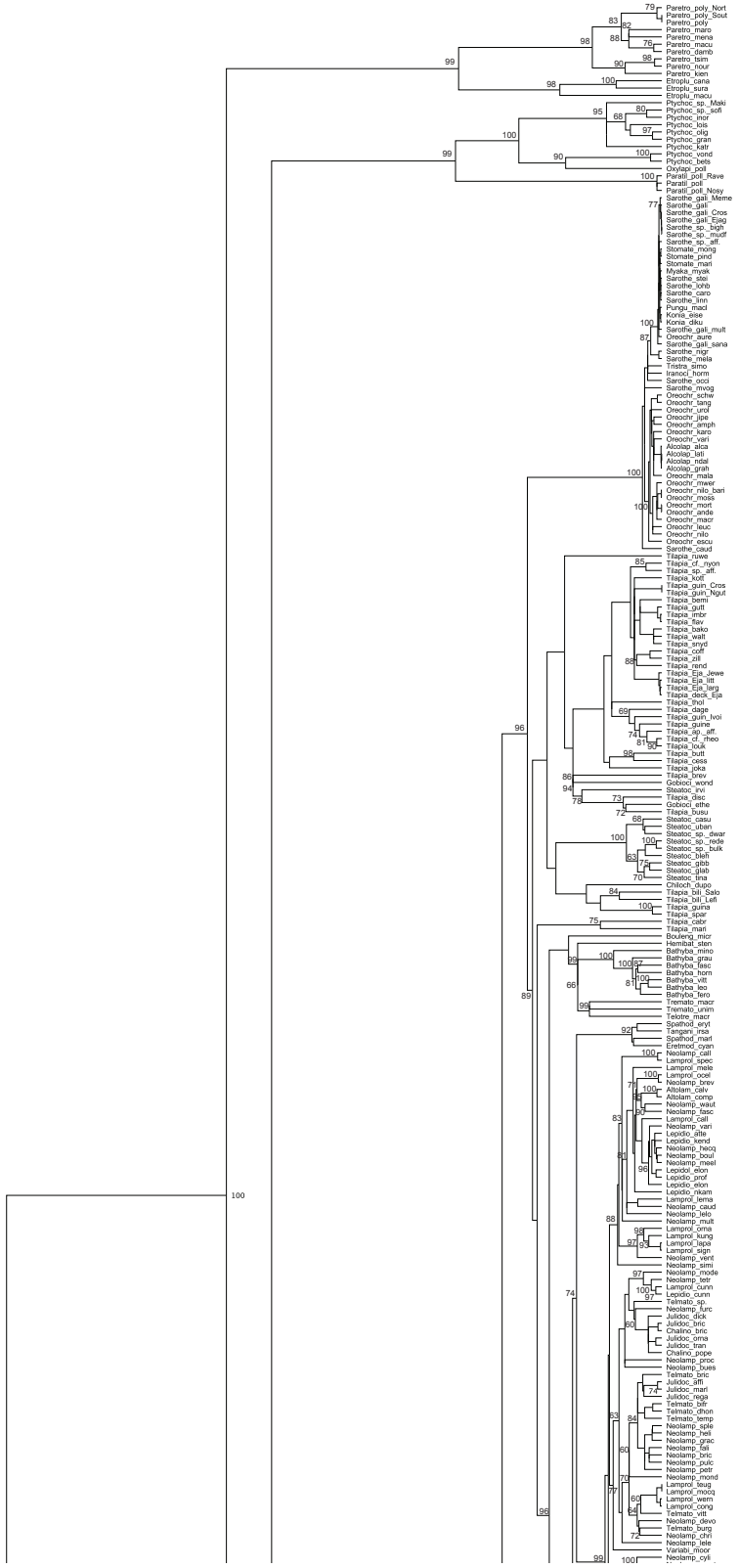
Supplementary Table 7. For radiations of five or more species, single predictor variable phylogenetic logistic regression reveals significant associations between radiation and lake depth, energy, latitude, mouthbrooding, time for diversification, and lake age. When Lake Tanganyika is excluded lake depth, egg spots, haplochromine egg spots, sexual dichromatism, time for diversification and lake age are significantly associated with radiation. N = sample size. a = phylogenetic signal in the diversification state/factor regression.

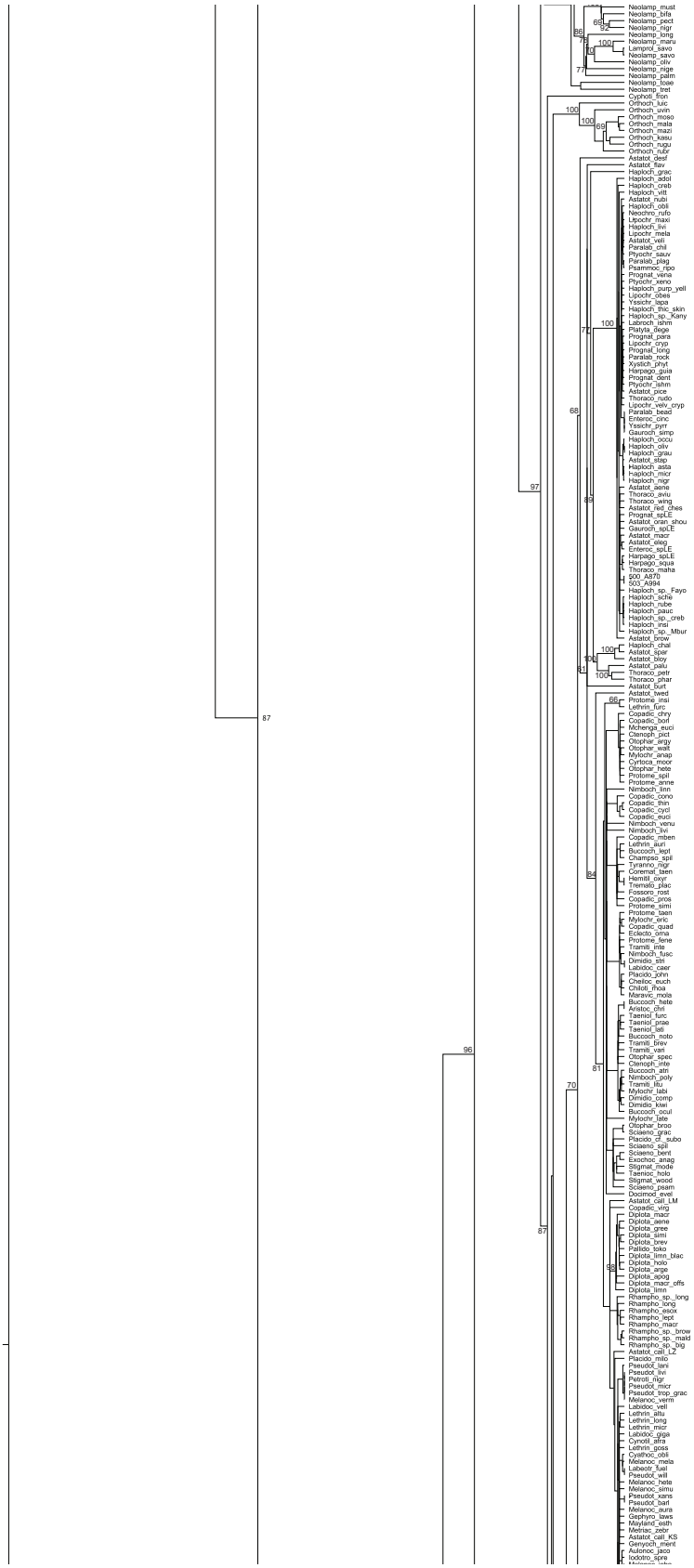
Predictor	Full Dataset						Excluding Tanganyika					
Environmental Variables	N	Estimate	SE (±)	z-value	z-value p	*a*	N	Estimate	SE (±)	z-value	z-value p	*a*
Lake Surface Area	166	0.088	0.054	1.642	0.050	-1.902	150	0.005	0.055	0.091	0.464	-2.699
Lake Depth	156	0.471	0.115	4.095	**0.000**	-2.451	140	0.363	0.153	2.367	**0.009**	-2.679
Energy	166	0.018	0.011	1.666	**0.048**	-2.477	150	0.014	0.014	1.070	0.142	-2.880
Latitude	166	-0.054	0.031	-1.738	**0.041**	-2.149	150	-0.052	0.041	-1.256	0.105	-2.786
Latitude (residual)	166	0.005	0.028	0.172	0.432	-2.065	150	0.004	0.044	0.092	0.463	-2.713
Elevation	166	0.001	0.000	1.355	0.088	-2.473	150	0.001	0.001	1.240	0.107	-2.813
Predators	158	-0.033	0.239	-0.139	0.445	-2.026	142	-0.524	0.459	-1.142	0.127	-2.697
Traits												
Polygamous Mating System	161	0.305	0.655	0.466	0.321	-2.150	148	0.975	0.851	1.146	0.126	-2.587
Mouthbrooding	166	1.923	0.905	2.125	**0.017**	-1.052	151	0.833	0.845	0.986	0.162	-2.645
Egg spots	166	0.428	0.780	0.548	0.292	-1.401	151	1.374	0.757	1.814	**0.035**	-2.667
Haplochromine egg spots	166	-0.759	0.936	-0.811	0.209	-1.414	151	1.456	0.694	2.097	**0.018**	-2.914
Sexual Dichromatism	161	0.187	0.432	0.433	0.332	-1.914	147	1.310	0.679	1.929	**0.027**	-2.537
Time												
Time in Lake	126	0.466	0.107	4.375	**0.000**	-2.487	113	0.521	0.197	2.640	**0.004**	-2.643
Lake Age	132	0.272	0.078	3.469	**0.000**	-2.424	117	0.194	0.116	1.670	**0.048**	-2.698

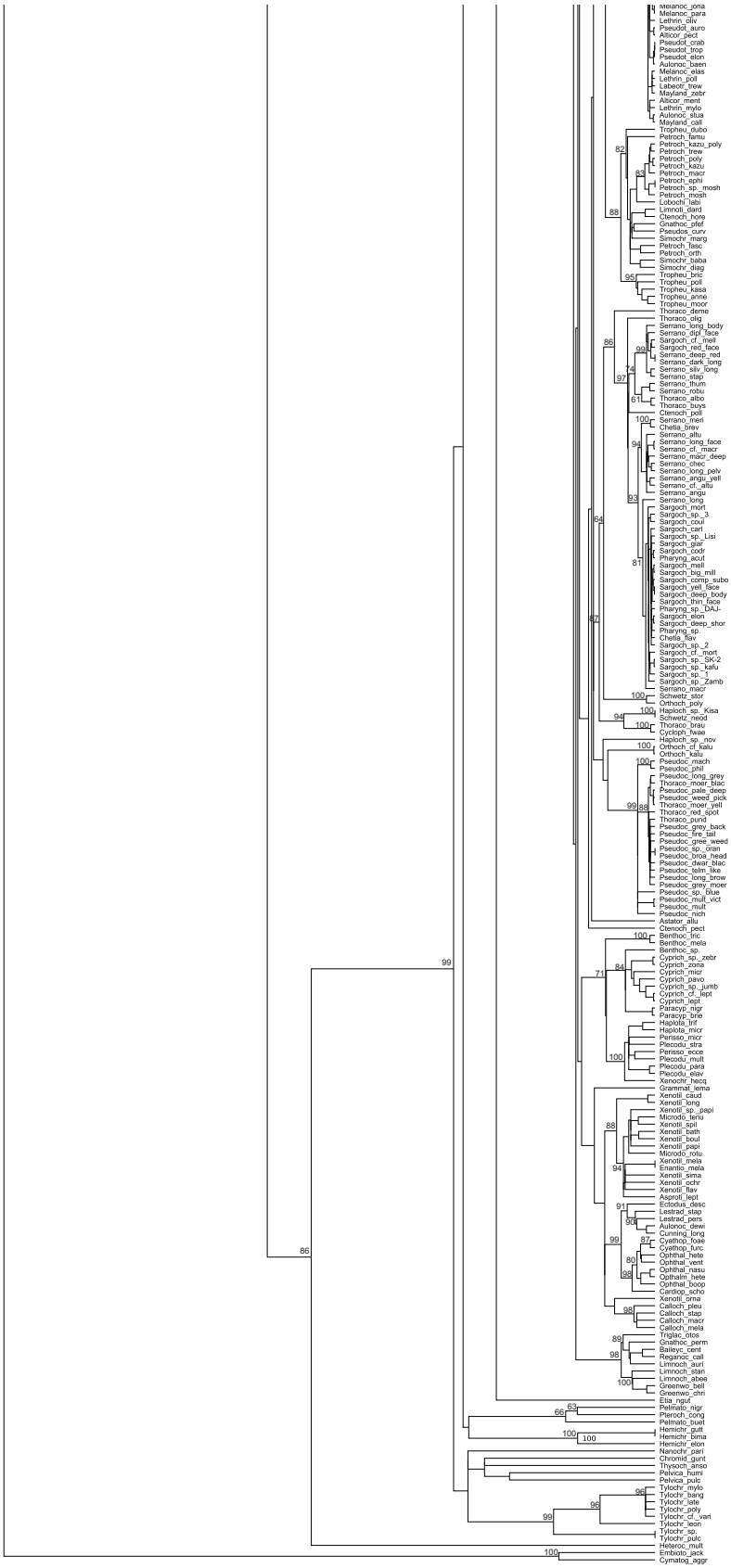
Supplementary Table 8. Multiple logistic regression models for diversification threshold 5. Lake depth and environmental energy remain the strongest environmental predictors of diversification, but lake surface area is no longer a strong predictor at this threshold. Sexual dichromatism remains the strongest intrinsic predictor of cichlid diversification.

	Full Dataset							Excluding Lake Tanganyika						
	nonphylogenetic			phylogenetic				nonphylogenetic			phylogenetic			
Predictor	Relative-importance value	Estimate	± SE	Estimate	±SE	Wald Z		Relative-importance value	Estimate	± SE	Estimate	±SE	Wald Z	
Lake Surface Area	0.484	-0.089	0.057					0.483	-0.095	0.061				
Lake Depth	**1.000**	0.636	0.170	0.438	0.118	3.711	**0.000**	0.610	0.243	0.137				
Energy	**0.748**	0.040	0.021	0.006	0.015	0.379	0.352	**0.860**	0.060	0.026	0.042	0.023	1.795	**0.036**
Residual Latitude	0.265	-0.007	0.023					0.263	-0.007	0.025				
Elevation	0.328	-0.000	0.000					0.606	-0.001	0.001				
Predators	0.339	0.205	0.223					0.333	-0.201	0.265				
Polygamous Mating System	0.358	-0.324	0.338					0.277	0.114	0.409				
Egg dummies	0.324	0.119	0.268					0.333	0.244	0.388				
Haplo egg dummies	0.629	0.821	0.461					**0.824**	1.666	0.743	1.257	0.736	1.707	**0.044**
Sexual Dichromatism	**0.708**	1.049	0.536	0.994	0.553	1.797	**0.036**	**0.708**	1.421	0.764	1.444	0.820	1.761	**0.039**

Supplementary Figure 1. Full tree inferred from phylogenetic analyses of the 653 taxon dataset. The topology here is the single best maximum likelihood tree; support on nodes are bootstrap support values from 100 rounds of bootstrapping in RAxML. (full tree is split over the next three pages)




Neolamp_must
Neolamp_bifa
Neolamp_pect
Neolamp_nigr
Neolamp_long
Neolamp_maru
Lamprol_savo
Neolamp_savo
Neolamp_oliv
Neolamp_nige
Neolamp_palm
Neolamp_toae
Neolamp_tret
Cyphoti_fron
Orthoch_luic
Orthoch_uvin
Orthoch_moso
Orthoch_mala
Orthoch_mazi
Orthoch_kasu
Orthoch_rugu
Orthoch_rubr
Astatot_desf
Astatot_flav
Haploch_grac
Haploch_adsd
Haploch_crab
Haploch_vitl
Astatot_nubi
Haploch_obli
Neochro_rufo
Lipochr_maxi
Haploch_livi
Lipochr_mela
Astatot_velt
Paralab_chil
Ptyochr_sauv
Paralab_plag
Psammoc_ripo
Prognat_vena
Ptyochr_veno
Haploch_purp_yell
Lipochr_obes
Yssichr_lapa
Haploch_thic_skin
Haploch_sp._Kany
Labroch_ishm
Platyla_ilege
Prognat_para
Lipochr_cryp
Prognat_long
Paralab_rock
Xystich_phyt
Herpego_guia
Prognat_dent
Ptyochr_lehm
Astatot_pice
Thoraco_rudo
Lipochr_velv_cryp
Paralab_bead
Enteroc_cinc
Yssichr_pyrr
Gauroch_simp
Haploch_occu
Haploch_oliv
Haploch_grau
Astatot_stap
Haploch_asta
haploch_micr
Haploch_nigr
Astatot_aene
Thoraco_aviu
Thoraco_wing
Astatot_red_ches
Prognat_spLE
Astatot_oran_shou
Gauroch_spLE
Astatot_macr
Astatot_aleg
Enteroc_spLE
Harpago_spLE
Herpego_aqua
Thoraco_maha
500_A870
503_A994
Haploch_sp._Fayo
Haploch_schie
Haploch_rube
Haploch_pauc
Haploch_sp._creb
Haploch_insi
Haploch_sp._Mbur
Astatot_brow
Haploch_chal
Astatot_spar
Astatot_bloy
Astatot_palu
Thoraco_petr
Thoraco_phar
Astatot_burl
Astatot_twed
Protome_insi
Lethrin_furc
Copadic_chry
Copadic_borl
Mchenga_euci
Ctenoph_pict
Otophar_argy
Otophar_wait
Mylochr_anap
Cyrtoca_moor
Otophar_hete
Protome_spil
Protome_anne
Nimboch_linn
Copadic_cono
Copadic_thin
Copadic_cycl
Copadic_euci
Nimboch_venu
Nimboch_livi
Copadic_mben
Lethrin_auri
Buccoch_lept
Champso_spil
Tyranno_nigr
Coremat_taen
Hemitil_oxyr
Trematо_plac
Fossoro_rost
Copadic_pros
Protome_simi
Protome_taen
Mylochr_eric
Copadic_quad
Eolecto_ornа
Protome_ferw
Tramiti_inte
Nimboch_fusc
Dimidio_stri
Labidoc_caer
Placido_john
Cheiloc_euch
Chiloti_rhoa
Maravic_mola
Buccoch_hete
Aristoc_chri
Taeniol_furc
Taeniol_prae
Taeniol_lati
Buccoch_noto
Tramiti_brev
Tramiti_vari
Otophar_spec
Ctenoph_inte
Buccoch_atri
Nimboch_poly
Tramiti_illu
Mylochr_labi
Dimidio_comp
Dimidio_kwi
Buccoch_ocul
Mylochr_late
Otophar_broo
Sciaeno_grac
Placido_cf._subo
Sciaeno_spli
Sciaeno_bent
Exochoc_anag
Stigmat_rhode
Taenioc_holo
Stigmat_wood
Sciaeno_psam
Docimod_evel
Astatot_call_LM
Copadic_virg
Diplota_mazr
Diplota_aene
Diplota_gree
Diplota_simi
Diplota_brev
Pallido_toko
Diplota_limn_blac
Diplota_holo
Diplota_arge
Diplota_apog
Diplota_mazr_offs
Diplota_limn
Rhampho_sp._long
Rhampho_long
Rhampho_esox
Rhampho_lepi
Rhampho_macr
Rhampho_sp._brow
Rhampho_sp._mald
Rhampho_sp._big
Astatot_call_LZ
Placido_mito
Pseudot_lani
Pseudot_livi
Petroti_nigr
Pseudot_micr
Pseudot_trop_grac
Melanoc_vermi
Labidoc_vell
Lethrin_altu
Lethrin_long
Lethrin_mizr
Labidoc_giga
Cynotil_afra
Lethrin_goss
Cyathoc_obli
Melanoc_meta
Labeotr_fuel
Pseudot_will
Melanoc_hete
Melanoc_simu
Pseudot_xans
Pseudot_bart
Melanoc_aura
Gephyro_laws
Mayland_esth
Melrac_zebr
Astatot_call_KS
Genyoch_mient
Aulonoc_jaco
Iodotro_spre

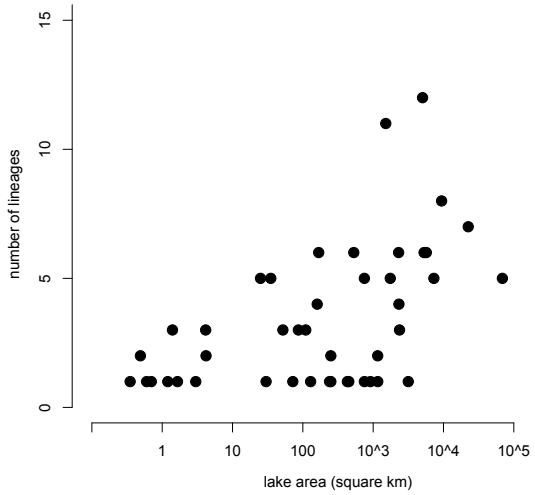

Melanoc_jona
Melanoc_para
Lethrin_oliv
Pseudot_auro
Alticor_pict
Pseudot_crab
Pseudot_trop
Pseudot_elon
Aulonoc_baen
Melanoc_elas
Lethrin_poll
Labeotr_trew
Mayland_zebr
Alticor_ment
Lethrin_mylo
Aulonoc_stua
Mayland_call
Tropheu_dubo
Petroch_famu
Petroch_kazu_poly
Petroch_trew
Petroch_poly
Petroch_kazu
Petroch_macr
Petroch_ephi
Petroch_sp._mosh
Petroch_mosh
Lobochi_labi
Limnoti_dard
Ctenoch_hore
Gnathoc_pfef
Pseudos_curv
Simochr_marg
Petroch_fasc
Petroch_orth
Simochr_baba
Simochr_diag
Tropheu_bric
Tropheu_poll
Tropheu_kasa
Tropheu_anne
Tropheu_moor
Thoraco_deme
Thoraco_ellg
Serrano_long_body
Serrano_dipl_face
Sargoch_cf._mell
Sargoch_red_face
Serrano_deep_red
Serrano_dark_long
Serrano_silv_long
Serrano_stap
Serrano_fhum
Serrano_robu
Thoraco_albo
Thoraco_buys
Ctenoch_poll
Serrano_meri
Chetia_brev
Serrano_altu
Serrano_long_face
Serrano_cf._macr
Serrano_macr_deep
Serrano_chec
Serrano_long_pelv
Serrano_angu_yell
Serrano_cf._altu
Serrano_angu
Serrano_long
Sargoch_mort
Sargoch_sp._3
Sargoch_coul
Sargoch_carl
Sargoch_sp._Lisi
Sargoch_glar
Sargoch_codr
Pharyng_acut
Sargoch_mell
Sargoch_big_mill
Sargoch_comp_subo
Sargoch_yell_face
Sargoch_deep_body
Sargoch_thin_face
Pharyng_sp._DAJ-
Sargoch_elon
Sargoch_deep_shor
Pharyng_sp.
Chetia_flav
Sargoch_sp._2
Sargoch_cf._mort
Sargoch_sp._SK-2
Sargoch_sp._kafu
Sargoch_sp._1
Sargoch_sp._Zamb
Serrano_macr
Schwetz_stor
Orthoch_poly
Haploch_sp._Kisa
Schwetz_neod
Thoraco_brau
Cyclogh_fwae
Haploch_sp._nov
Orthoch_cf._kalu
Orthoch_kalu
Pseudoc_mach
Pseudoc_phil
Pseudoc_long_grey
Thoraco_moer_blac
Pseudoc_pale_deep
Pseudoc_weed_pick
Thoraco_moer_yell
Thoraco_red_spot
Thoraco_pund
Pseudoc_grey_back
Pseudoc_fire_tail
Pseudoc_gree_weed
Pseudoc_sp._oran
Pseudoc_broa_head
Pseudoc_dwar_blac
Pseudoc_telm_like
Pseudoc_long_brow
Pseudoc_grey_moer
Pseudoc_sp._Blue
Pseudoc_mult_vict
Pseudoc_mult
Pseudoc_nich
Astator_allu
Ctenoch_pect
Benthoc_lric
Benthoc_mela
Benthoc_sp.
Cyprich_sp._zebr
Cyprich_zona
Cyprich_micr
Cyprich_pavo
Cyprich_sp._jumb
Cyprich_cf._lepl
Cyprich_lepl
Paracyp_nigr
Paracyp_brie
Haplota_trif
Haplota_micr
Perisso_micr
Plecodu_stra
Perisso_ecce
Plecodu_mult
Plecodu_para
Plecodu_elav
Xenochr_hecq
Grammat_lema
Xenotil_caud
Xenotil_long
Xenotil_sp._papi
Microdo_tenu
Xenotil_spil
Xenotil_bath
Xenotil_boul
Xenotil_papi
Microdo_rotu
Xenotil_mela
Enantio_mela
Xenotil_sima
Xenotil_ochr
Xenotil_flav
Asproti_lept
Ectodus_desc
Lestrad_stap
Lestrad_pens
Aulonoc_dewi
Cunning_long
Cyathop_foae
Cyathop_furc
Ophthal_hete
Ophthal_vent
Ophthal_nasu
Opthalmi_hete
Ophthal_boop
Cardiop_scho
Xenotil_orna
Calloch_pleu
Calloch_stap
Calloch_macr
Calloch_mela
Triglac_otos
Gnathoc_perm
Baileyc_cent
Reganoc_call
Limnoch_auri
Limnoch_stan
Limnoch_abea
Greenwo_bell
Greenwo_chri
Etia_nguf
Petrtilo_nigr
Pteroch_cong
Petrtilo_buet
Hemichr_gutt
Hemichr_bima
Hemichr_elon
Nanochr_pari
Chromid_gunt
Thysoch_anso
Pelvica_humi
Pelvica_pulc
Tylochr_mylo
Tylochr_bang
Tylochr_late
Tylochr_poly
Tylochr_cf._vari
Tylochr_leon
Tylochr_sp.
Tylochr_pulc
Heteroc_mult
Embiotc_jack
Cymatog_aggr

APPENDIX 4

Supplementary Information for Chapter 4

1) Can a correlation between lake size and lineage sampling explain species-area patterns?

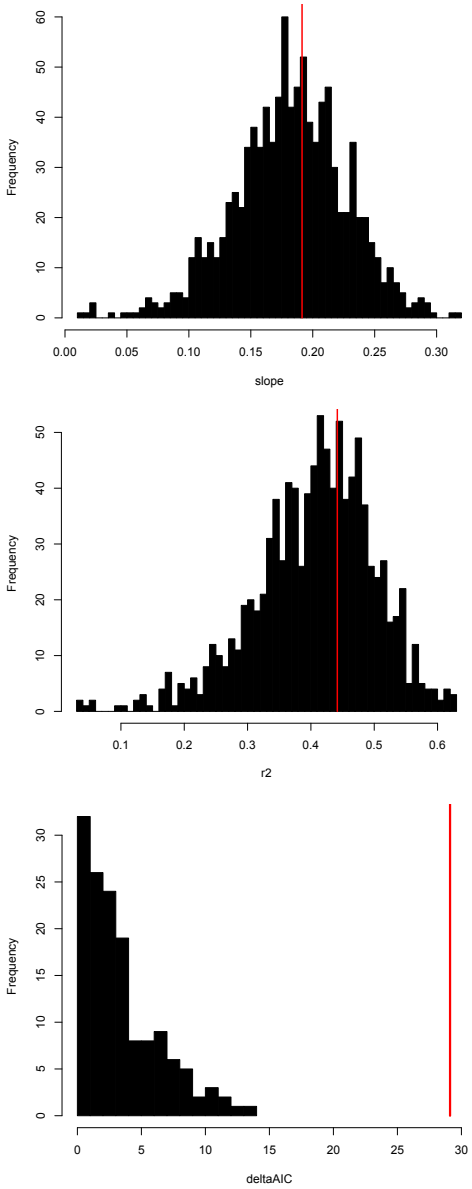
Lake size is likely at least roughly correlated with watershed area and/or the number of tributaries entering and exiting the lake. Therefore, we might expect that as lake size increases, the number of cichlid lineages present in the lake would increase simply as a result of increased chances of colonization with larger area. We used plots and quantile regression analyses to examine the relationship between lake size and the number of lineages present per lake. We find that size constrains the number of lineages found in a lake, but does not determine it (Supplementary Figure 1).



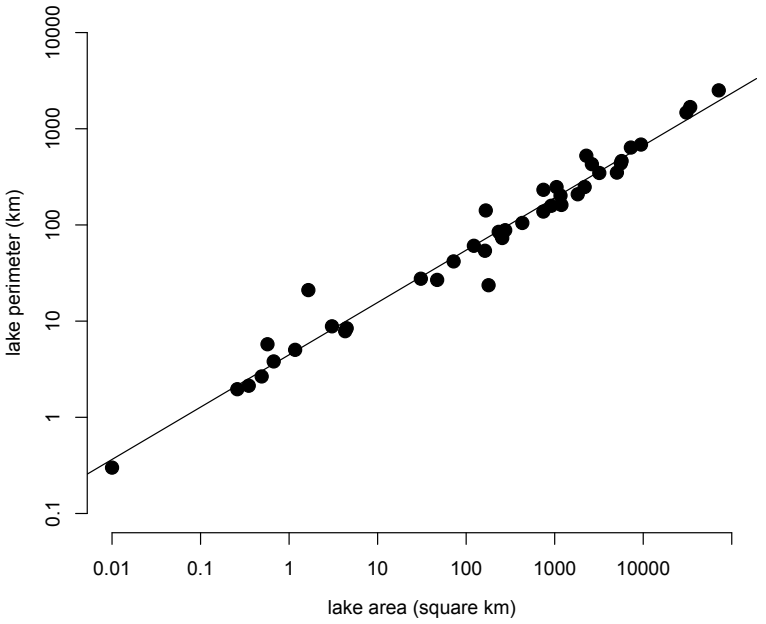
Supplementary Figure 1. The number of lineages in a lake is constrained but not determined by lake area.

To investigate whether the observed species-area pattern (Figure 4.1a) could result from the compounded effects of increased lineage sampling with increased area alone, we performed a permutation test. We retained the real number of cichlid lineages present in each lake, and we randomly drew the species richness for each of these lineages from the observed distribution of lineage species richnesses in our total dataset. We sampled this distribution without replacement for each lake, and then compared the slope and correlation of the linear species-area relationship from 1000 resampled datasets to our observed data. To assess whether a two-slope regression relationship could result from lineage sampling effects alone, we calculated ΔAIC values comparing linear and broken regression models for the resampled data and compared these to our observed value.

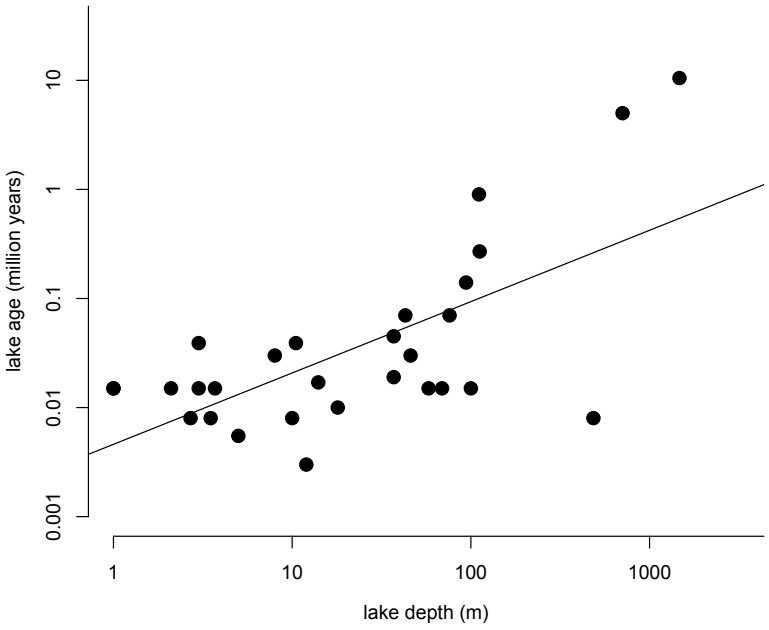
Results of these permutations show that, we can produce linear model slopes and correlations well within those that we find for the actual data, but we cannot not reproduce the observed breakpoint relationship (Supplementary Figure 2). Therefore, the observed breakpoint relationship cannot result from a lake size-derived increase in the number of colonizing lineages alone.



Supplementary Figure 2. Permuting the richness of lineages within lakes reveals that the (a) slope and (b) correlation we observe for a linear relationship between species richness and area is well within the distribution of values that we see for the resampled datasets (red line = observed value). However, only in <35% of the dataset allow us to evaluate a broken regression model, and of these only half support a broken regression model over a linear model, and (c) within this subset none come anywhere near the support for the broken regression model over the linear regression model that we see in the real dataset (assessed by ΔAIC).



Supplementary Figure 3. There is a very strong positive correlation between lake area and perimeter ($r^2 = 0.97$)



Supplementary Figure 4. There is a significant positive correlation between lake depth and lake age ($r^2 = 0.43$)

Supplementary Table 1. Lake ages or time since last desiccation, whichever is more recent. Where age ranges are given in the source, numbers here are the midpoint of age ranges. Starred numbers represent ages inferred from molecular phylogenies of cichlid lineages within the lake.

Lake	Age/Time since Desiccation	References
Abaya	--	
Abbe	0.019	Barker & Gasse 2003
Albert	0.015	Beuning et al. 1997
Bangweulu	0.008	Scholz et al. 2007
Baringo	0.015	Owen et al. 1982; Garcin et al. 2009; Bergner et al. 2009
Barombi ba Koto	--	
Barombi Mbo	0.900 *	see this volume, Chapter 3
Bermin	0.750 *	Schliewen et al. 1994
Bosumtwi	0.070	Scholz et al. 2007
Chad	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Chala	0.140	Moernaut et al. 2010
Chamo	--	
Chilwa	0.008	Scholz et al. 2007
Chiuta	0.008	Scholz et al. 2007
Debo	0.039	Cohen et al. 2007; Street-Perrott & Perrott 1993
Edward	0.270	Beuning et al. 1997; Laerdal & Talbot 2002
Ejagham	0.010	Schliewen et al. 2001
Eyasi	0.015	Johnson et al. 2000; Scholz et al. 2007 ; Stager & Johnson 2008
Fwa	--	
Guinas sink hole	--	
Jipe	--	
Kagera lakes	--	
Kinneret	0.070	Hazan et al. 2005
Kivu	0.008	Haberyan & Hecky, 1987
Lake Mai Ndombe	--	
Lake Tschungruru/Masoko	0.050	Barker & Gasse 2003
Langano	0.030	Street 1979; Barker & Gasse 2003
Logipi	--	
Lutoto	0.050	Sato et al. 2003
Malawi	5.000	Scholz & Finney 1994
Manyara	0.015	Johnson et al. 2000; Scholz et al. 2007 ; Stager & Johnson 2008
Mareotis	--	
Mweru	0.045	Scholz et al. 2007
Nabugabo	0.006	Stager et al. 2005
Natron	0.015	Barker & Gasse 2003
Nshere	0.050	Sato et al. 2003
Rukwa	0.015	Barker & Gasse 2003
Saka	0.003	Bessems et al., 2003; Russell et al., 2007
Stefani	--	
Tana	0.017	Lamb et al. 2007
Tanganyika	10.500	Cohen et al. 1997
Tumba	--	
Turkana	0.015	Owen et al. 1982; Beadle 1974; Garcin et al. 2009; Bergner et al. 2009
Upemba lakes	--	
Victoria	0.015	Johnson et al. 2000
Zwai	0.030	Street 1979; Barker & Gasse 2003

Supplementary Table 2. Equations used for testing among nested multivariate models incorporating lake area, lake depth, and energy. For results see Table 1.

Model	Equation
Area	$\log(S) = c + z1 * \log(A)$
Depth	$\log(S) = c + z1 * \log(D)$
Total Energy	$\log(S) = c + z1 * E_{tot}$
Depth + Energy	$\log(S) = c + z1 * \log(D) + z2 * E_{tot}$
Area + Energy	$\log(S) = c + z1 * \log(A) + z2 * E_{tot}$
Area + Depth	$\log(S) = c + z1 * \log(A) + z2 * \log(D)$
Area + Depth + Energy	$\log(S) = c + z1 * \log(A) + z2 * \log(D) + z3 * E_{tot}$
Two-slope Area	$\log(S) = c + z1 * \log(A) + z2[(\log(A)) - t) * (\log(A) \geq t)]$
Two-slope Area + Depth	$\log(S) = c + z1 * \log(A) + z2[(\log(A)) - t) * (\log(A) \geq t)] + z3 * \log(D)$
Two-slope Area + Energy	$\log(S) = c + z1 * \log(A) + z2[(\log(A)) - t) * (\log(A) \geq t)] + z3 * E_{tot}$
Two-slope Area + Energy + Depth	$\log(S) = c + z1 * \log(A) + z2[(\log(A)) - t) * (\log(A) \geq t)] + z3 * \log(D) + z4 * E_{tot}$