5-2014

Adaptive radiation along a benthic/pelagic ecological axis in North America's most diverse, endemic clade of freshwater fishes

Phillip Ray Hollingsworth Jr.
University of Tennessee, Knoxville, phollin1@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Phillip Ray Hollingsworth Jr. entitled "Adaptive radiation along a benthic/pelagic ecological axis in North America's most diverse, endemic clade of freshwater fishes." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Ecology and Evolutionary Biology.

Christopher D. Hulsey, Major Professor

We have read this dissertation and recommend its acceptance:

James A. Fordyce, Brian C. O'Meara, Richard J. Strange

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Adaptive radiation along a benthic/pelagic ecological axis in North America's most diverse, endemic clade of freshwater fishes

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Phillip Ray Hollingsworth Jr.
May 2014

DEDICATION

I dedicate this dissertation to my late father Phillip Ray Hollingsworth Sr. and my late grandfather Robert Henderson Hollingsworth Jr. These men influenced my life and this work immeasurably.

ACKNOWLEDGEMENTS

I would first like to thank my academic advisor and friend Dr. C. Darrin Hulsey for his guidance throughout my graduate studies, as well as for introducing me to world fish diversity during a number of memorable field expeditions. I would also like to thank the members of my dissertation committee, Dr. Brian O'Meara, Dr. Richard Strange, and Dr. James Fordyce for their academic guidance. I am particularly thankful to Dr. Fordyce and Dr. O'Meara for their instruction concerning the statistical and computational methodology employed in this dissertation. I owe a great amount of gratitude to Dr. David Etnier who greatly influenced the development of my abilities as a naturalist and as an educator in freshwater biology. I would like to thank my past and current lab mates for social support and academically productive conversation. Finally, I would like to thank the Department of Ecology and Evolutionary Biology at UTK, the Alexander Hollaender Fellowship, and the Cokkinias Graduate Fellowship for funding this work.

ABSTRACT

Eastern North America is unparalleled throughout the temperate world in terms of freshwater fish biodiversity. A monophyletic group of approximately 250 cyprinid fishes, known as the open posterior myodome (OPM) clade, dominates the fish species richness in the freshwater ecosystems of this region. In this dissertation, I explore the influence of eco-evolutionary divergence along a benthic/pelagic habitat axis on the generation of this hyper-diverse group of fishes. My three chapters work synergistically to address the question: Did a historical shift from benthic to pelagic habitats by OPM cyprinids represent the invasion of an open adaptive zone and result in the simultaneous bursts of phylogenetic and ecological diversification that signify an adaptive radiation? In Chapter I, I perform the first gene tree/species tree analysis on OPM species to reconcile discordance between previous phylogenetic hypotheses as it relates to inferring the history of benthic and pelagic habitat transitions in the group. I then construct the most thoroughly sampled OPM phylogenies to date in Chapter II. Using these large-scale phylogenies and habitat-use data, I conducted ancestral state reconstructions to trace the history of benthic to pelagic habitat use during the history of the clade. I then performed lineage through time and diversification rate analyses that suggested that a period of accelerated lineage diversification followed the initial shift from benthic to pelagic habitats in the OPM radiation. In Chapter III, I recovered a significant evolutionary relationship between jaw protrusion angle (JPA) and preferred foraging height in the water column between 15 co-occurring OPM taxa. I also recovered evidence for a burst of morphological disparification in a number of individual muscoskeletal characters that are evolutionary correlated with JPA after the major benthic to pelagic shift inferred in Chapter II. Overall, the results from this dissertation suggest that an early shift from benthic to pelagic habitats in OPM cyprinids represented the invasion of an open adaptive zone and was followed by a period of rapid phylogenetic and eco-morphological evolution as species diverged to exploit vertically segregated sub-zones throughout the water column. Taken together, these results are likely the most robust evidence for an adaptive radiation to date.

TABLE OF CONTENTS

INTRODUCTION 1
References 2
CHAPTER I RECONCILING GENE TREES OF EASTERN NORTH AMERICAN MINNOWS 4
Abstract 5
Introduction 5
Methods 7
Results 11
Discussion 13
References 17
CHAPTER II EXPLOSIVE DIVERSIFICATION FOLLOWING A BENTHIC TO PELAGIC SHIFT IN FRESHWATER FISHES 20
Abstract 21
Introduction 21
Methods 23
Results 26
Discussion 28
References 31
CHAPTER III ADAPTIVE RADIATION ALONG A BENTHIC/PELAGIC AXIS IN NORTH AMERICA'S MOST DIVERSE, ENDEMIC CLADE OF FRESHWATER FISHES 36
Abstract 37
Introduction 37
Methods 40
Results 43
Discussion 45
References 49
CONCLUSION 54
APPENDIX 55
VITA 73

LIST OF TABLES

Table 1.1 Genetic partition information..... 56
Table 2.1 Tests for an early burst of diversification in the pelagic clade 57
Table 3.1 Preferred foraging height and JPA across the Little River OPM assemblage .. 58
Table 3.2 Correlation coefficients, r, of PICs between JPA and habitat use and JPA and individual morphological characters..... 59
Table 3.3 Comparison of 2-Rate versus 1-Rate Brownian motion models of trait evolution before and after the pelagic transition..... 60
Table 3.4 Comparison of 'Early Burst' versus Brownian motion models of trait evolution after the pelagic transition..... 61

LIST OF FIGURES

Figure 1.1 Individual gene trees 62
Figure 1.2 Multi-locus phylogenies 63
Figure 2.1 Concatenated Cyt*b* + Rag1 phylogeny 64
Figure 2.2 Ancestral state reconstructions 67
Figure 2.3 Diversification rate analyses for the focal clade 68
Figure 2.4 Simulated lineage accumulation and ΔAIC values 69
Figure 3.1 Mean JPA and phylogenetic relationships based on Rag1 across the Little River OPM community 70
Figure 3.2 Morphological measurements utilized in Chapter III 71
Figure 3.3 Relative evolutionary rate decrease after the pelagic shift and towards the present for four morphological correlates of JPA 72

LIST OF ATTACHMENTS

File 1.1 Specimen information
File 1.2 Sequencing primers
File 1.3 GenBank accession numbers
File 2.1 GenBank numbers and habitat designations
File 2.2 Individual MCC gene trees

INTRODUCTION

Freshwater fishes often diversify along a benthic (bottom) to pelagic (mid-water) habitat axis and experience shifts in functional morphology associated with these two habitat types (Robinson and Wilson, 1994; Willacker et al., 2010). Robinson and Wilson (1994), citing evidence from across a wide range of fish phylogenetic diversity, argue for a predictable pattern of evolution into preexisting benthic and pelagic "niches" within lacustrine environments. Often, this morphological divergence is considered evidence for the process of ecological speciation (Schluter, 2000). However, these studies have mainly addressed this pattern in relatively depauparate, lentic systems at high latitudes (e.g., Bodalay, 1979; Skúlason et al., 1989; Schluter and McPhail, 1992; Schluter, 1993; Robinson et al., 1996; Svanbäck and Eklöv, 2003). Further, the goal of most of these studies was to identify intraspecific divergence into two "niches" within competitor poor communities (Robinson and Wilson, 1994). It remains unclear if this predictable pattern of microevolutionary divergence between benthic and pelagic forms plays a significant macroevolutionary role in the diversification and community assembly of species-rich freshwater fish clades inhabiting more ecologically complex, lotic freshwater environments.

Cyprinid fishes from eastern North America provide an attractive study system for pursuing questions regarding freshwater fish evolution along a benthic/pelagic habitat axis in lotic environments. With approximately 250 species native to this region, cyprinids comprise a significant portion the species diversity within the most diverse assemblages of temperate freshwater fishes in the world. Previous phylogenetic analyses suggest that the vast majority of these 250 species represent a monophyletic group united by the osteological synapomorphy known as the open posterior myodome (OPM) (Mayden, 1989; Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010; Hollingsworth et al., 2013). OPM cyprinids have diversified extensively along a benthic/pelagic axis and form complex communities consisting of up to 15 sympatric taxa within many river systems in the southeastern U.S. (Etnier and Starnes, 1993; Boschung and Mayden, 2004; Page and Burr, 2011). Coexistence in these diverse assemblages is facilitated by the partitioning of the water column into vertically segregated microhabitats from the benthos to the surface (Mendelson, 1975; Baker and Ross, 1981; Gorman, 1988a,b)

This dissertation researches the diversification of OPM cyprinids along a benthic/pelagic habitat axis. I constructed the most thoroughly sampled phylogenetic hypotheses for the clade to date and combined them with data on habitat-use and trophic morphology to address the question: Did a major benthic to pelagic shift early in the history of the group represent the invasion of a novel adaptive zone that was followed by a period of rapid phylogenetic and eco-morphological diversification suggestive of an adaptive radiation? The results from this dissertation should elucidate the role of macroevolution along a benthic/pelagic habitat axis in the generation of North America's most speciose, endemic clade of freshwater fishes.

References

Baker, J.A., Ross, S.T., 1981. Spatial and temporal resource utilization by southeastern cyprinids. *Copeia* 1981, 178-189.

Boschung, H.T., Mayden, R.L., 2004. Fishes of Alabama. Smithsonian Books, Washington, D.C.

Bodalay, R.A., 1979. Morphological and ecological divergence within the lake whitefish (*Coregonus clupeaformis)* species complex in Yukon Territory. *J. Fish. Res. Board Can.* 36, 1214-1222.

Bufalino, A.P., Mayden, R.L., 2010. Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. *Mol. Phylogenet. Evol.* 55, 274-283.

Etnier, D.A., Starnes, W.C., 1993. The fishes of Tennessee. University of Tennessee Press, Knoxville, TN.

Gorman, O.T., 1988a. The dynamics of habitat use in a guild of Ozark minnows. *Ecol. Monogr.* 58, 1-18.

Gorman, O.T., 1988b. An experimental study of habitat use in an assemblage of Ozark minnows. *Ecology* 69, 1239-1250.

Hollingsworth Jr., P.R., Simons, A.M., Fordyce, J.A., Hulsey, C.D., 2013. Explosive diversification following a benthic to pelagic shift in freshwater fishes. *BMC Evol. Biol.* 13, 272.

Houston, D.D., Shiozawa, D.K., Riddle, B.R., 2010. Phylogenetic relationships of the western North American cyprinid genus *Richardsonius*, with an overview of phylogeographic structure. *Mol. Phylogenet. Evol.* 55, 259-273.

Mayden, R.L., 1989. Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: Cypriniformes). *Misc. Pub. Univ. Kans. Mus. Nat. Hist.* 80, 1-189.

Mendelson, J., 1975. Feeding relationships among species of *Notropis* (Pisces: Cyprinidae) in a Wisconsin stream. *Ecol. Monogr.* 45, 199-230.

Page, L.M., Burr, B.M., 2011. Peterson field guide to freshwater fishes, second edition. Houghton Mifflin, Boston, MA.

Robinson, B.W., Wilson, D.S., 1994. Character release and displacement in fishes: a neglected literature. *Am. Nat.* 144, 596-627.

Robinson, B.W., Wilson, D.S., Shea, G.O., 1996. Tradeoffs of ecological specialization: an intraspecific comparison of pumpkinseed sunfish phenotypes. *Ecology* 77, 170-178.

Schluter, D., 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology* 74, 699–709.

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

Schluter, D., McPhail, J.D., 1992. Ecological character displacement and speciation in sticklebacks. *Am. Nat.* 140, 85-105.

Simons, A.M., Berendzen, P.B., Mayden, R.L., 2003. Molecular systematics of North American phoxinin genera (Actinopterygii : Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. *Zool. J. Linn. Soc. Lond*. 139, 63-80.

Skúlason, S., Noakes, D.L.G., Snorrason, S.S., 1989. Ontogeny of trophic morphology in four sympatric morphs of arctic charr *Salvelinus alpinus* in Thingvallavatn, Iceland. *Biol. J. Linn. Soc.* 38, 281–301.

Svanbäck, R., Eklöv, P., 2003. Morphology dependent foraging efficiency in perch: a trade-off for ecological specialization? *Oikos* 102, 273-284.

Willacker, J.J., von Hippel, F.A., Wilton, P.R., Walton, K.M., 2010. Classification of threespine stickleback along the benthic-limnetic axis. *Biol. J. Linn. Soc.* 101, 595-608.

CHAPTER I
RECONCILING GENE TREES OF EASTERN NORTH AMERICAN MINNOWS

A version of this chapter was published by Phillip R. Hollingsworth Jr. and C. Darrin Hulsey. PRH conceived the study, conducted the analyses, and wrote the manuscript. CDH helped to prepare the manuscript:

Hollingsworth Jr., P.R., Hulsey, C.D., 2011. Reconciling gene trees of eastern North American minnows. *Molecular Phylogenetics and Evolution* 61, 149-156.

Abstract

Most eastern North American cyprinid fishes belong to a clade known as the "open posterior myodome" (OPM) minnows, but phylogenetic relationships within this clade have been difficult to ascertain. Previous attempts to resolve relationships among the generally benthic "chubs" and the more pelagic "shiners", that constitute the majority of OPM minnows, have led to highly discordant phylogenetic hypotheses. To further examine relationships among the OPM minnows, we utilized both a concatenated Bayesian approach and a coalescent-based species tree method to analyze data from six protein coding nuclear loci (Enc1, Ptr, Ryr3, Sh3px3, Tbr1, and Zic1), as well as the mitochondrial locus (Cytb). We focused our analyses on the chub-like genus *Phenacobius,* a group that has drifted topologically between other benthic chubs and the more pelagic shiners, and also included exemplar taxa from 11 other OPM lineages. Individual gene trees were highly discordant regarding relationships within *Phenacobius* and across the OPM clade. The concatenated Bayesian analysis and coalescent-based species tree reconstruction recovered slightly different phylogenetic topologies. Additionally, the posterior support values for clades using the coalescent-based approach were consistently lower than the concatenated analysis. However, *Phenacobius* was resolved as monophyletic and as the sister lineage to *Erimystax* regardless of the combined data approach taken. Furthermore, *Phenacobius* + *Erimystax* was recovered as more closely related to the shiners we examined than to other chubs. Relationships within *Phenacobius* varied depending on the combined phylogenetic method utilized. Our results highlight the importance of multi-locus, coalescent-based approaches for resolving the phylogeny of diverse clades like the eastern North American OPM minnows.

Introduction

Fishes in the family Cyprinidae dominate the freshwater habitats of North America (NA) with over 300 species distributed from Canada south to the Neovolcanic Plateau in southern Mexico (Burr and Mayden, 1992). The eastern half of NA contains a particularly high number of cyprinid lineages, and the vast majority comprise a single clade united by the morphological synapomorphy of an open posterior myodome (OPM) (Mayden, 1989; Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010). It has been hypothesized based on osteological characters that there are two major monophyletic groups within the hyperdiverse OPM clade: 1) the generally benthic "chubs" and 2) the more pelagic "shiners" (Mayden, 1989). However,

in subsequent phylogenetic studies, relationships among constituent chub and shiner lineages have varied considerably (Coburn and Cavender, 1992; Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). In order to clarify the relationships among major lineages of minnows in eastern NA, we analyzed sequence data of six nuclear DNA (nDNA) loci and one mitochondrial DNA (mtDNA) locus using exemplars of recognized genera within the OPM clade. We also included complete taxon sampling of species in the genus *Phenacobius* in order to examine relationships within this genus as well as the relationship of *Phenacobius* to other OPM genera. *Phenacobius* was placed in the chub clade by Mayden (1989) and Coburn and Cavender (1992) based on osteology, but recent molecular phylogenies suggest it may be more closely related to the shiner OPM minnows (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Therefore, determining the phylogenetic position of this genus of five species poses an interesting problem and focal point for this multi-locus phylogenetic analysis of OPM minnow relationships.

Phylogenetic trees reconstructed from individual loci do not necessarily reflect the species tree of a given clade because discordance among gene trees may be common, especially in clades characterized by large population sizes and short intervals between diversification events (Degnan and Rosenberg, 2009). Such discordance among gene trees and the desire to infer the best species tree for a group challenges our prior understanding of how phylogenetic relationships should be examined (Maddison, 1997; Maddison and Knowles, 2006; Degnan and Rosenberg, 2006, 2009). For instance, simulations show that simply adding more data to a concatenated matrix can lead to faulty inferences concerning the true species tree when incomplete lineage sorting, or deep coalescence, causes discordance between phylogenetic markers (Kubatko and Degnan, 2007). To account for this shortcoming, methods have been developed that take phylogenetic information from multiple loci and model their evolution under multi-species coalescent theory in an attempt to account for incomplete lineage sorting (Edwards et al., 2007; Kubatko et al., 2009; Liu et al., 2009; Heled and Drummond, 2010). These methods are now being implemented across a broad range of vertebrate groups, but have largely focused on closely related species (White et al., 2009; Carstens and Dewey, 2010; McCormack et al., 2011). Incomplete lineage sorting of gene trees could also be a problem deeper in the phylogenetic history of a clade of organisms, especially if the timeframe for gene coalescence is greater than the period during which groups became genetically isolated. As previous analyses of OPM minnow relationships have had difficulties resolving relationships not only among recently diverged groups but also among genera deep in the clade, variability in gene coalescence may be problematic to phylogenetic inference at several levels in this group.

To evaluate if the stochasticity of coalescence poses a problem for estimating the phylogeny among closely related OPM minnow species, groups such as *Phenacobius* would be interesting to examine within a gene tree/species tree phylogenetic framework. All five named *Phenacobius* species are morphologically diagnosable and generally allopatrically distributed, leaving little question of species monophyly (Etnier and Starnes, 1993; Jenkins and Burkhead, 1994). This lack of sympatry among members of

Phenacobius should also lead to relatively little hybridization among these five species which is a potentially complicating source of incongruence among gene trees that is not accounted for in most current gene tree/species tree methods (Degnan and Rosenberg, 2009). Additionally, all previous analyses of OPM phylogeny included *Phenacobius* and the placement of this group has varied substantially (Mayden, 1989; Coburn and Cavender, 1992; Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Finally, Dimmick and Burr (1999) conducted a study of the phylogenetic relationships among the five species of *Phenacobius* based on a combination of morphological, allozyme, and DNA sequence data which provides a hypothesis with which to compare the results from the molecular phylogenetic approaches taken in this study.

Most previous studies of OPM phylogenetics have relied heavily on mitochondrial and nuclear intron sequence data (Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006). Only recently has sequence data from protein coding nuclear loci been utilized in phylogenetic analyses of NA minnows (Bufalino and Mayden, 2010; Schonhuth and Mayden, 2010; Hulsey and Hollingsworth, 2011). Therefore, a major goal of this study was to generate sequence data for six putatively single copy, protein-coding nuclear loci to serve as phylogenetic markers within NA minnows. In this study, we combine sampling among most OPM genera with complete taxon sampling within the genus *Phenacobius* in order to contrast results from a coalescent-based species tree phylogenetic approach with a concatenated Bayesian analysis of protein-coding nDNA at various levels of phylogenetic relationships.

We used a recently developed multi-species coalescent-based phylogenetic strategy implemented through *BEAST (Heled and Drummond 2010) to examine the relationships among 16 species of OPM minnows, and to compare to results from a concatenated Bayesian analysis. We utilized sequences from seven loci to address a number of questions concerning these relationships. First, we asked what the phylogenetic relationships among exemplars of several of the most diverse OPM genera are. Then, we asked how the genus *Phenacobius* is related phylogenetically to the other OPM minnow genera analyzed in this study. Finally, we examined the phylogenetic relationships among the five species of *Phenacobius*. For each of the above questions, we compared the results between the concatenated Bayesian analysis and the species tree approach implemented in *BEAST (Heled and Drummond, 2010).

Methods

DNA sequence generation

Specimens sequenced in this study were collected in the field using a seine net. Locality information and museum accession numbers are given in the attachments (File 1.1). Specimens of all five species in the OPM genus *Phenacobius* were included, as well as representative taxa from eleven eastern NA OPM genera. These eleven taxa included species designated by Mayden (1989) as chubs (*Campostoma oligolepis*, *Erimystax dissimilis*, *Exoglossum laurae*, and *Nocomis effusus*), as well as taxa designated by

Mayden (1989) as shiners (*Cyprinella callistia*, *Luxilus coccogenis*, *Lythrurus fasciolaris*, and *Notropis leuciodus*). We also included the species *Pimephales notatus* and *Hybopsis amblops* that were hypothesized by Mayden (1989) to fall outside of the chub and shiner clades, but that have been recovered as nested within the shiner clade in subsequent phylogenetic analyses (Coburn and Cavender, 1992; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). *Rhinichthys cataractae* was also included. Morphological phylogenies (Mayden, 1989; Coburn and Cavender, 1992) did not include *Rhinichthys* spp. in the OPM clade. However, more recent molecular phylogenies have included *Rhinichthys* as an early diverging lineage within the OPM radiation and closely related to other chub genera (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Overall, the taxa included here span the phylogenetic breadth of previously proposed hypotheses of OPM relationships. Individual specimens were anesthetized in MS-222 prior to removal of a pectoral fin for a tissue sample. Tissue samples were stored in 1.5 mL tubes in 95% EtOH and placed in an -80° C freezer for long-term storage. DNA extraction was performed using a Qiagen DNeasy kit (Qiagen Sciences, MD, USA).

PCR amplification was carried out using an Eppendorf DNA thermocycler. The sequences for all primer sets used in this study are given in the attachments (File 1.2). The mitochondrial cytochrome *b* (Cytb) gene was amplified using primers from Schmidt and Gold (1993) for all species except *Phenacobius mirabilis*, *P. teretulus*, and *P. uranops*. Cytb was amplified for these three species using the primer set MinCytb F2 and MinCytb R1. The six nuclear loci examined included: ectodermal-neural cortex 1 (Enc1), hypothetical protein LOC 564097 (Ptr), novel protein similar to vertebrate ryanodine receptor 3 (Ryr3), protein similar to SH3 and PX domain containing 3 gene (Sh3px3), T-box brain 1 (Tbr1), and zic family member 1 (Zic1). All six nuclear genes were sequenced using primers from Li et al. (2007). PCR conditions consisted of an initial denaturation phase at 94° C (2 min) followed by 35 cycles of 94° C (1 min), 54° C (1 min), and 72° C (1 min). A final elongation phase of 72° C (4 min) was performed after the cycles in order to ensure complete elongation of amplified products.

DNA sequencing was performed at the University of Washington's High Throughput Genomics Unit utilizing the same primers used during PCR. Sequence files were contiged using Sequencher 4.8 (Gene Codes, Ann Arbor, MI, USA) and heterozygous sites in the nuclear loci were coded as ambiguous using the IUPAC codes for heterozygous sites. In addition to the newly generated sequences generated in this study, sequence data for Cytb and Enc1 from Hulsey and Hollingsworth (2011) was downloaded from GenBank for several species of NA cyprinids for use in phylogenetic analyses. Sequence data for all loci examined were downloaded from GenBank for the Asian species *Danio rerio* which was used as an outgroup in all phylogenetic analyses. Additionally, the non-OPM NA minnow *Semotilus atromaculatus* was included as an outgroup to the OPM minnows. Due to difficulties in PCR amplification, the Cytb sequence data for *Semotilus atromaculatu*s was also downloaded from GenBank (Dowling et al., 2002). GenBank accession numbers for all sequence data are given in the attachments (File 1.3). Sequences were aligned using Clustal X (Thompson et al., 1997). Codon sites were defined using MacClade 4.0 (Maddison and Maddison, 2000).

Gene tree reconstruction

We first inferred the gene tree for each individual locus in order to compare the topologies among the individual genes. Because of the number of informative sites in the mitochondrial gene, each codon position for Cytb was assigned a separate model of molecular evolution. The nuclear loci were not partitioned into individual codon sites because the low variability in their first and second codon positions provided little information to estimate parameters for a separate model of molecular evolution. The first codon position of Zic1 was monomorphic across the North American species analyzed in this study and was therefore excluded in the phylogenetic analyses. The best model of molecular evolution for each locus was chosen using MrModelTest2 (Nylander, 2004). MrModelTest2 starts with a neighbor-joining tree for each partition and then calculates likelihood scores for each of the 24 substitution models that can be implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). The best substitution model for each genetic partition was then chosen based on the model with the lowest AIC score, which penalizes models consisting of more free parameters. Models chosen for each locus are presented along with the length in base pairs of the sequences examined at each locus (Table 1.1). All tables and figures are given in the Appendix at the end of this dissertation.

Model parameters were then designated in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) in order to approximate the maximum likelihood tree for each locus. For the Cytb analyses, the command prset ratepr=variable was used to allow for rate variation between codon positions. The gamma shape distribution, proportion of invariant sites, state frequencies, and relative rates of substitution were estimated separately for the three codon positions of Cytb using the unlink command in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Individual MCMC analyses of each locus consisted of two independent runs of four chains and were run for 1,000,000 generations with trees and parameter estimates sampled every 100 generations. The default heating temperature of 0.1 was used in these analyses. Each MCMC analysis was run three separate times.

Convergence in the MrBayes analyses was assessed by analyzing the split frequencies between the two simultaneous but independent MCMC runs using the "compare" and "cumulative" plots in the program AWTY (Nylander et al., 2008). The program Tracer (Drummond and Rambaut, 2007) was also used to assess convergence by monitoring likelihood and ESS, or effective sample size, values through the course of each MCMC run. The ESS is a proxy for the amount of mixing of Markov chains and represents the number of independent draws from the posterior distribution. High ESS values (> 200) signify sufficient mixing of Markov chains, and consequently, low amounts of autocorrelation between parameter estimates from one generation to the next during the course of the MCMC run. Based on all convergence diagnostics, each run had converged by 50,000 generations and the first 100,000 generations for each MCMC search were discarded as the burn-in. Then, the remaining post burn-in trees were used to construct a 50% majority rule consensus tree for each individual locus using the sumt command. Posterior probability values were averaged across the three independent MCMC runs for each locus.

Multi-locus species tree reconstructions

Two methods were utilized in order to reconstruct the species trees from the combination of the individual loci. For both analyses, we first concatenated the data into a matrix containing 5393 characters. For the concatenated analysis, we used MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) to generate a concatenated phylogeny using all 7 loci. In this analysis, we specified 9 partitions corresponding to the three codon positions of Cytb and the six unpartitioned nuclear loci (minus Zic1 first position sites). The same models of molecular evolution that were applied in individual gene tree reconstructions were assigned to the nine partitions (Table 1.1). The ratepr=variable command was applied in order to allow for rate variation across partitions, and the gamma shape distribution, proportion of invariant sites, state frequencies, and relative rates of substitution were unlinked across partitions. Each concatenated MrBayes run consisted of two separate runs of four chains for 10,000,000 generations with a sampling of trees and parameter values every 1,000 generations, and heating temperature of 0.1. Convergence was assessed both in AWTY (Nylander, 2008) and using scale reduction factors reported from MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Tracer (Drummond and Rambaut, 2007) was also used to graphically depict likelihood and ESS values over the course of the runs. Because all runs appeared to have converged by 500,000 generations, the first 1,000,000 runs were subsequently discarded as the burn-in period. We ran three independent MCMC searches for this dataset and averaged the posterior probability values for the nodes across the three replicates to produce our consensus concatenated phylogeny.

The program *BEAST (Heled and Drummond, 2010) was also used to estimate a species tree from the individual loci. First, the partitioned alignment of 5393 characters from the concatenated analysis was imported into the program BEAUti v1.5.4 (Drummond and Rambaut, 2007). We then unlinked substitution models, clock models, and trees across each partition, with the exception of Cytb in which the trees were linked between the three codon positions. The same substitution models used in the concatenated analysis were assigned to each partition (Table 1.1). A relaxed molecular clock for each partition was estimated relative to Enc1 with all rate estimates drawn from an uncorrelated lognormal distribution. The ploidy level of the Cytb partition was designated as "mitochondrial" and all nuclear loci were designated "autosomal nuclear". Each *BEAST (Heled and Drummond, 2010) search consisted of 1.0×10^8 generations, sampling trees every 1000 generations and used the default priors from BEAUti v1.5.4 (Drummond and Rambaut, 2007). We ran five replicate species tree searches using *BEAST (Heled and Drummond, 2010). Convergence was assessed after importing the log file from each run into Tracer (Drummond and Rambaut, 2007) and then monitoring likelihood and ESS values through the course of the run. To ensure the burn-in was sufficient and to allow our computer to efficiently run the program, the first 9.0×10^7 generations were discarded from each run and then the post burn-in trees and parameter estimates were combined from the five independent runs to produce a majority-rule consensus tree using LogCombiner (Drummond and Rambaut, 2007).

Results

Patterns of sequence variation and individual gene trees

Maximum uncorrected sequence divergence between North American species utilized in this study and the two species displaying the divergence are provided for each partition (Table 1.1). The most variable partition was Cytb third codon position (52.5%) and the least was Tbr1 (2.0%). The divergence reported for the Zic1 locus (Table 3) does not include first position codon sites, as this nucleotide position was invariable. The non-OPM creek chub, *Semotilus atromaculatus*, was one of the species involved in the maximum observed sequence divergence in six of the nine partitions with the exception of the Cytb second position, Sh3px3, and Zic1 (Table 1.1).

The Bayesian 50% majority-rule consensus tree for each individual locus is depicted (Figure 1.1). Relationships among the included genera of OPM minnows varied between individual gene trees. The non-OPM taxon *Semotilus atromaculatus* was resolved as falling outside of, or within a polytomy that was sister to, all other NA lineages sampled in five of the seven individual gene trees. In the Sh3px3 gene tree, *Semotilus atromaculatus* was recovered as sister to *Nocomis effusus* with low posterior support (56%). In the Ptr phylogeny, *Semotilus atromaculatus* and *Nocomis effusus* were recovered as more closely related to *Erimystax* + *Phenacobius* spp. than to other NA lineages sampled. This relationship received moderate posterior support (81%) in this gene tree.

The remaining chub lineages, *Campostoma oligolepis, Exoglossum laurae*, and *Rhinichthys cataractae*, were resolved as diverging early in the majority of the individual gene trees (4 of 7) along with *Nocomis effusus* (Figure 1.1). However, relationships between these taxa and the remainder of the OPM taxa included in this study varied considerably, often receiving low posterior support, depending on the individual gene tree. For example, *Exoglossum laurae* is strongly supported (91% posterior) as diverging after *Nocomis effusus*, but before the remainder of the OPM minnows in the Cytb gene tree. However, in the Enc1 gene tree *Exoglossum laurae* and *Rhinichthys cataractae* are moderately supported as sister taxa (85% posterior). Relationships between the shiner taxa included in this study were also variable. Only two sets of relationships between shiner taxa were resolved in more than one individual gene tree. The sister relationship between *Pimephales notatus* and *Cyprinella callistia* was recovered with 100% and 89% posterior support in the Ryr3 and Enc1 gene trees respectively. *Luxilus coccogenis* and *Notropis leuciodus* were also recovered as each other's closest relative with 93% and 65% posterior support in the Zic1 and Cytb gene trees respectively (Figure 1.1).

Erimystax dissimilis was recovered as sister to *Phenacobius* spp. with high posterior support (>90%) in all individual gene trees except for the generally poorly resolved Tbr1 gene tree (Figure 1.1). This clade of *Erimystax* + *Phenacobius* was recovered as closely aligned with the NA shiners (*Cyprinella callistia*, *Hybopsis amblops*, *Lythrurus fasciolaris*, *Luxilus coccogenis*, and *Notropis leuciodus*) in four of the seven individual gene trees with substantial posterior Bayesian support (>95%). Relationships between *Erimystax* + *Phenacobius* and the other OPM lineages sampled

were poorly resolved in the Ptr and Tbr1 gene trees. However, in the Enc1 gene tree *Erimystax* + *Phenacobius* was recovered in 53% of post burn-in trees as sister to a moderately supported clade (88% posterior) of all remaining OPM minnows (Figure 1.1).

Phenacobius was significantly supported as monophyletic in 6 of 7 gene trees (Figure 1.1). However, there was little agreement among the gene trees concerning the phylogenetic relationships within *Phenacobius*, and many, but not all, of these relationships received low posterior support values. The most consistently resolved node within *Phenacobius* was a sister relationship between *P. uranops* and *P. crassilabrum*. This relationship was recovered in 87%, 97%, and 100% of the post burn-in Ryr3, Seh3px3, and Cytb gene trees, respectively. Relationships between the three other members of *Phenacobius* varied considerably across gene trees. *Phenacobius catostomus* and *P. mirabilis* were recovered as sister lineages with moderate posterior support, 70% and 79%, in the Cytb and Ryr3 gene trees (Figure 1.1). The two species *P. teretulus* and *P. catostomus* were recovered as sister species in the Sh3px3 and Zic1 gene trees with 86% and 99% posterior probability values respectively.

Multi-locus species trees

In the concatenated analysis, *Semotilus atromaculatus* was recovered as the outgroup to a strongly supported (100% posterior) clade consisting of the remaining NA lineages sampled (Figure 1.2A). All but four nodes in the concatenated topology received significant posterior support of > 95%. These remaining ambiguous nodes subtended the following sets of taxa: 1) *Phenacobius mirabilis* and *P. uranops* (54%), 2) *P. catostomus* and *P. teretulus* (82%), 3) *Hybopsis amblops* and *Rhinichthys cataractae* (89%), and 4) *Exoglossum laurae* and *Rhinichthys cataractae* (88%). The chub genera were recovered at the base of the concatenated OPM phylogeny with *Nocomis effusus* strongly supported (100% posterior) as the earliest diverging lineage of the OPM minnows analyzed. This divergence was followed by divergence of other chub genera, first *Campostoma* and then a clade of *Exoglossum* + *Rhinichthys*, with moderate to strong posterior support (88-100%) (Figure 1.2A). Within the shiners, *Lythrurus fasciolaris* was recovered as the earliest diverging lineage. *Hybopsis amblops* and *Luxilus coccogenis* formed a strongly supported clade, while *Notropis leuciodus* was recovered as sister to a clade of *Pimephales notatus* + *Cyprinella callistia*. All relationships within the shiners received high (98-100%) Bayesian posterior support in the concatenated analysis (Figure 1.2A).

Phenacobius was strongly supported (100% posterior) as monophyletic and *Erimystax dissimilis* was strongly supported as sister to *Phenacobius* spp. (100% posterior) in the concatenated phylogeny (Figure 1.2A). The clade containing both *Erimystax* + *Phenacobius* received 100% posterior support as being sister to the shiners included in this analysis. Within *Phenacobius*, a weakly supported (54% posterior) clade of *P. mirabilis* and *P. uranops* + *P. crassilabrum* was recovered as sister to *P. catostomus* + *P. teretulus* (Figure 1.2A).

The topology estimated by *BEAST (Heled and Drummond, 2010) is very similar to the concatenated topology (Figure 1.2B). However, posterior support values were noticeably lower in this coalescent-based phylogenetic reconstruction. Therefore, all

support values including those below 50% on the *BEAST topology are presented as maximum clade credibility scores in order to compare with results from the concatenated topology. Using the *BEAST (Heled and Drummond, 2010) species tree approach, the chubs are recovered as diverging first from the remaining OPM lineages followed by divergence of *Erimystax* + *Phenacobius* from the shiner clade (Figure 1.2B). However, the species tree analysis differed from the concatenated analysis in specific relationships both among the chubs and shiners (Figure 1.2). *Nocomis effusus* was again recovered as the earliest diverging OPM lineage, albeit with low posterior support (51%). However, *Campostoma oligolepis* was strongly supported (94% posterior) as sister to the *Exoglossum* + *Rhinichthys* clade contrary to the concatenated analysis. Within the shiners, *Lythrurus fasciolaris* was recovered as sister to *Notropis leuciodus* with low posterior support (37%). Among the shiners, only the *Pimephales notatus* + *Cyprinella callistia* clade resolved in the concatenated analysis was again recovered with high posterior support (91%) by the *BEAST (Heled and Drummond, 2010) species tree analysis (Figure 1.2B).

Erimystax dissimilis was strongly supported (100% posterior) as sister to a strongly supported (100% posterior) monophyletic *Phenacobius* in the coalescent-based species tree (Figure 1.2B). This clade was recovered in 100% of post burn-in species trees as more closely related to the shiner taxa utilized. However, within *Phenacobius,* a different set of relationships was resolved by *BEAST (Heled and Drummond, 2010) than those recovered in the concatenated analysis (Figure 1.2). Although the two species *P. uranops* and *P. crassilabrum* were again recovered as sister lineages, this clade was recovered as sister to a clade of *P. catostomus* + *P. teretulus* in 46% of the post burn-in species trees, with *P. mirabilis* resolved as the earliest diverging member of *Phenacobius* (Figure 1.2B).

Discussion

The individual gene trees of eastern NA OPM minnows were highly discordant. The two species tree approaches also produced phylogenetic hypotheses that differed topologically in several areas. Additionally, our concatenated phylogeny recovered much higher posterior support values as compared to those obtained from the coalescent-based species tree strategy. Despite incongruence between the two multi-locus approaches, both methods of reconstructing the species tree provided a generally concordant backbone topology. Both phylogenies indicated that OPM shiners are monophyletic and nested within the OPM clade, making chubs a paraphyletic group. Both species trees also resolved *Phenacobius* + *Erimystax* as more closely related to the shiner taxa than to the chub taxa that we analyzed. The topologies of the two multi-locus trees are likely to have been strongly influenced by the phylogenetic signal within the Cytb locus due to its greater variability relative to the nuclear loci at this phylogenetic level. In the future, it will be interesting to see if sequencing many more nuclear genes provides a topology that is consistently divergent from that of Cytb. We will also be able to determine if these nuclear genes converge on a singular set of relationships for regions of the OPM minnow phylogeny that are currently difficult to resolve.

One area of discordance between the two multi-locus analyses involved relationships among the chub genera. For example, relationships among the early diverging chub lineages were strongly supported in the concatenated species tree, with *Nocomis* recovered as the earliest diverging OPM taxa. In the coalescent species tree, this sister relationship between *Nocomis* and the remainder of the OPM minnows was not well supported. Additionally, *Campostoma* was recovered as the next chub lineage to diverge after *Nocomis* from the remaining OPM lineages in the concatenated analysis. However, in the coalescent approach *Campostoma* is supported as sister to a clade of *Exoglossum* + *Rhinichthys.* These sets of relationships among chub lineages differed from all previous, large-scale molecular phylogenies of OPM minnows (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011), in which *Nocomis* and *Campostoma* were hypothesized to be sister lineages. Interestingly, the sister group relationship between *Exoglossum* and *Rhinichthys* that was supported by both of our species tree analyses was also recovered by Simons et al. (2003). Mayden et al. (2006) and Hulsey and Hollingsworth (2011) did not include *Exoglossum* species in their analyses, but this relationship was not recovered in the phylogeny presented by Bufalino and Mayden (2010).

Relationships and posterior support values also differed substantially between our two multi-locus approaches within the shiner clade. The concatenated analysis produced a fully resolved, and well-supported, topology of shiner genera relationships. However, the coalescent-based species tree generally provided little support for shiner intergeneric relationships. The only significantly supported node within the shiner clade that we recovered using both species tree strategies was the sister relationship between *Pimephales* and *Cyprinella*. Hulsey and Hollingsworth (2011) presented a topology containing *Pimephales* nested within an unsupported, yet monophyletic *Cyprinella.* However, this relationship was not recovered by any of the other aforementioned studies of OPM molecular phylogenies (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010).

The lower posterior probability values on the coalescent-based species tree can be partially explained by the more parameter-rich phylogenetic model implemented by this strategy. The increased number of parameters in the coalescent model results from taking into consideration the discordance between individual gene trees. Coalescent methods must also estimate parameters affecting coalescent processes such as ancestral population sizes (Heled and Drummond, 2010). Concatenation of data ignores this discordance between gene trees, estimates less parameters, and can lead to inflated posterior support values (Kubatko and Degnan, 2007). The discordance we recovered between our individual gene trees suggest that accounting for the coalescent process is likely warranted when examining relationships among OPM genera and at multiple other levels of phylogenetic inference.

Despite incongruence and variable support between our two multi-locus phylogenetic approaches in several areas of the phylogeny, both analyses recovered the strongly supported, sister relationship between *Phenacobius* + *Erimystax* and the shiner clade. This relationship is consistent with other recent molecular phylogenies produced for this group of fishes (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden,

2010; Hulsey and Hollingsworth, 2011). However, earlier morphological and molecular phylogenies proposed that *Erimystax* and *Phenacobius* were more closely related to the remaining chub OPM lineages sampled in this study (Mayden, 1989; Coburn and Cavendar, 1992; Simons and Mayden, 1997). *Erimystax* and *Phenacobius* are benthic taxa and generally chub-like in appearance. As such, our species tree analyses suggest that chub-like morphology and its association with benthic habitats were common early in the history of the eastern NA OPM clade and persisted at least until the divergence of *Phenacobius* + *Erimystax* from the remaining eastern OPM minnows. Interestingly, the number of species in these early diverging chub lineages is relatively low in comparison to extant shiner diversity. The pelagic shiners likely arose from ancestral benthic forms and subsequently diversified to generate a substantial portion of contemporary OPM diversity. Evolution along a benthic/limnetic habitat continuum has been demonstrated to be a common axis of diversification between closely related species in several lacustrine fish groups (Hatfield and Schluter, 1999; Barluenga et al., 2006; Bertrand et al. 2008). This study suggests that this benthic/limnetic axis of diversification could also be an important macroevolutionary force in generating species diversity across broader phylogenetic scales within lotic environments such as the streams of eastern NA.

There was little agreement between gene trees and species trees in resolving the phylogenetic relationships within *Phenacobius*. However, both of our multi-locus phylogenetic analyses resolved different relationships than those posited by Dimmick and Burr (1999). Their analysis based on morphological and genetic data recovered *Phenacobius mirabilis* as the earliest diverging lineage within *Phenacobius*, followed in order of divergence by *P. teretulus*, then by *P. catostomus*, and finally by a clade of *P. uranops* + *P. crassilabrum*. Both our concatenated and species-tree analyses also supported the sister species relationship between *P. uranops* and *P. crassilabrum*, the two species with abutting allopatric ranges in the upper Tennessee River system. However, contrary to the hypothesis of Dimmick and Burr (1999), the species *P. teretulus* and *P. catostomus* were weakly supported as each other's closest relative in both of our species tree approaches. Furthermore, our concatenated analysis recovered a weakly supported sister relationship between *P. mirabilis* and the clade of *P. crassilabrum* + *P. uranops*. A lack of phylogenetically informative variation within *Phenacobius* at the nDNA loci utilized could partially explain the incongruence between gene trees and uncertainty in our species trees.

Several of the same regions of the phylogeny that we had trouble resolving in this study (relationships between chubs lineages and within the shiners and *Phenacobius*) have differed topologically in previous molecular phylogenetic studies of OPM relationships (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Furthermore, these regions have consistently been recovered as areas with particularly short branch lengths between diverging lineages (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). These are regions of the phylogeny that have likely experienced large amounts of incomplete lineage sorting of slowly evolving genetic markers due to short time frames between diversification events. Future studies of OPM minnow phylogenetics should aim at generating sequence data from faster evolving nuclear loci,

and employ coalescent-based species tree methods, in order to tease apart relationships in these problematic regions of the OPM phylogeny.

With the increasing availability of genomic data from across the tree of life, the ability to analyze DNA sequence data for multiple loci will no longer be limiting to phylogenetics. A transition from single locus to multi-locus, coalescent-based phylogenetics is well underway (Maddison and Knowles, 2006; Edwards et al., 2007; Degnan and Rosenberg, 2009; Kubatko et al., 2009; Liu et al., 2009; Heled and Drummond, 2010; Hulsey et al., 2011). Only by gathering data from independent genetic regions across the genomes of organisms, may we gain a better understanding of the phylogenetic relationships recorded in each locus. Individual gene phylogenies may then be analyzed separately under coalescent-based methods in order to approximate the distribution of species trees suggested by individual locus data. Future studies of phylogenetics of the hyper-diverse North American OPM minnows should give less weight to phylogenetic hypotheses based on single genetic partitions and concatenated multi-locus matrices and place more emphasis on producing phylogenies that explicitly model the coalescent process.

References

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

Bertrand, M., Marcogliese, D.J., Magnan, P., 2008. Trophic polymorphism in brook charr revealed by diet, parasites and morphometrics. *J. Fish Biol.* 72, 555-572.

Bufalino, A.P., Mayden, R.L., 2010. Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. *Mol. Phylogenet. Evol.* 55, 274-283.

Burr, B.M., Mayden, R.L., 1992. Phylogenetics and North American freshwater fishes. In: Mayden, R.L. (Ed.), Systematics, historical ecology, and North American freshwater fishes. Stanford University Press, Stanford, CA, pp. 18-75.

Carstens, B.C., Dewey, T.A., 2010. Species delimitation using a combined coalescent and information-theoretic approach: an example from North American *Myotis* bats. *Syst. Biol.* 59, 400-414.

Coburn, M.M., Cavender, T.M., 1992. Interrelationships of North American freshwater fishes. In: Mayden, R.L. (Ed.), Systematics, historical ecology, and North American freshwater fishes. Stanford University Press, Stanford, CA, pp. 328-373.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet.* 2, e68.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332-340.

Dimmick, W.W., Burr, B.M., 1999. Phylogenetic relationships of the suckermouth minnows, genus *Phenacobius*, inferred from parsimony analyses of nucleotide sequence, allozymic and morphological data (Cyprinidae: Cypriniformes). Biochem. *Syst. and Ecol.* 27, 469-485.

Dowling, T.E., Tibbets, C.A., Minckley, W.L., Smith, G.R., 2002. Evolutionary relationships of the plagopterins (Teleostei: Cyprinidae) from cytochrome *b* sequences. *Copeia* 2002, 665-678.

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

Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. *P. Natl. Acad. Sci. USA* 104, 5936-5941.

Etnier, D.A., Starnes, W.C., 1993. The fishes of Tennessee. The Unviersity of Tennessee Press, Knoxville, TN.

Hatfield, T., Schluter, D., 1999. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. *Evolution* 53, 866-873.

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

Hulsey, C.D., Keck, B.P., Hollingsworth Jr., P.R., 2011. Species tree estimation and the historical biogeography of heronie cichlids. *Mol. Phylogenet. Evol.* 58, 124-131.

Hulsey, C.D., Hollingsworth Jr., P.R., 2011. Do constructional constraints influence cyprinid (Cyprinidae: Leuciscinae) craniofacial coevolution? *Biol. J. Linn. Soc*. 103, 136-146.

Jenkins, R.E., Burkhead, N.M., 1994. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, MD.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56, 17-24.

Kubatko, L.S., Carstens, B.C., Knowles, L.L., 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 25, 971-973.

Li, C.H., Orti, G., Zhang, G., Lu, G.Q., 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol. Biol.* 7, 44.

Liu, L., Yu, L., Kubatko, L., Pearl, D.K., Edwards, S.V., 2009. Coalescent methods for estimating phylogenetic trees. *Mol. Phylogenet. Evol.* 53, 320-328.

Maddison, W.P., 1997. Gene trees in species trees. *Syst. Biol.* 46, 523-536.

Maddison, D.R., Maddison, W.P., 2000. Macclade. Sinauer Associates, Inc., Sunderland, MA.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55, 21-30.

Mayden, R.L., 1989. Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: Cypriniformes). *Misc. Pub. Univ. Kans. Mus. Nat. Hist.* 80, 1-189.

Mayden, R.L., Simons, A.M., Wood, R.M., Harris, P.M., Kuhajda, B.R., 2006. Molecular systematics and classification of North American Notropin shiners and minnows (Cypriniformes: Cyprinidae). In: De Lourdes Lozano-Villano, M.A., Contreras-Balderas, A.J. (Eds.), Studies of North American desert fishes in honor of E.P. (Phil) Pister, conservationist. Universidad Autonoma de Nuevo Leon, Nuevo Leon, Mexico, pp. 72-101.

McCormack, J.E., Heled, J., Delaney, K.S., Townsend Peterson, A., Knowles, L.L., 2011. Calibrating divergence times on species trees versus gene trees: implication for speciation history of *Aphelocoma* jays. *Evolution* 65, 184-202.

Nylander, J.A.A., 2004. MrModeltest v2. Uppsala University.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* 24, 581-583.

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

Schmidt, T.R., Gold, J.R., 1993. Complete sequence of the mitochondrial cytochome-b gene in the cherryfin shiner, *Lythrurus roseipinnis* (Teleostei, Cyprinidae). *Copeia* 1993, 880-883.

Schonhuth, S., Mayden, R.L., 2010. Phylogenetic relationships of the genus *Cyprinella* based on mitochondrial and nuclear gene sequences. *Mol. Phylogenet. Evol.* 55, 77-98.

Simons, A.M., Mayden, R.L., 1997. Phylogenetic relationships of the creek chubs and the spine-fins: an enigmatic group of North American cyprinid fishes (Actinopterygii : Cyprinidae). *Cladistics* 13, 187-205.

Simons, A.M., Berendzen, P.B., Mayden, R.L., 2003. Molecular systematics of North American phoxinin genera (Actinopterygii : Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. *Zool. J. Linn Soc.-Lond.* 139, 63-80.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876-4882.

White, M.A., Ane, C., Dewey, C.N., Larget, B.R., Payseur, B.A., 2009. Fine-scale phylogenetic discordance across the house mouse genome. *PLoS Genet.* 5, e1000729.

CHAPTER II
EXPLOSIVE DIVERSIFICATION FOLLOWING A BENTHIC TO PELAGIC SHIFT IN FRESHWATER FISHES

A version of this chapter was published by Phillip R. Hollingsworth Jr., Andrew M. Simons, James A. Fordyce, and C. Darrin Hulsey. PRH conceived the study, conducted the analyses, and wrote the manuscript. AMS contributed sequence data. JAF assisted with statistical analyses. CDH helped to prepare the manuscript:

Hollingsworth Jr., P.R., Simons, A.M., Fordyce, J.A., Hulsey, C.D., 2013. Explosive diversification following a benthic to pelagic shift in freshwater fishes. *BMC Evolutionary Biology* 13, 272.

Abstract

Interspecific divergence along a benthic to pelagic habitat axis is ubiquitous in freshwater fishes inhabiting lentic environments. In this study, we examined the influence of this habitat axis on the macroevolution of a diverse, lotic radiation using mtDNA and nDNA phylogenies for eastern North America's most species-rich freshwater fish clade, the open posterior myodome (OPM) cyprinids. We used ancestral state reconstruction to identify the earliest benthic to pelagic transition in this group and generated fossil-calibrated estimates of when this shift occurred. This transition could have represented evolution into a novel adaptive zone, and therefore, we tested for a period of accelerated lineage accumulation after this historical habitat shift. Ancestral state reconstructions inferred a similar and concordant region of our mtDNA and nDNA based gene trees as representing the shift from benthic to pelagic habitats in the OPM clade. Two independent tests conducted on each gene tree suggested an increased diversification rate after this inferred habitat transition. Furthermore, lineage through time analyses indicated rapid early cladogenesis in the clade arising after the benthic to pelagic shift. A burst of diversification followed the earliest benthic to pelagic transition during the radiation of OPM cyprinids in eastern North America. As such, the benthic/pelagic habitat axis has likely influenced the generation of biodiversity across disparate freshwater ecosystems.

Introduction

Freshwater fish are frequently thought to diversify along a benthic (bottom) to pelagic (mid-water) habitat axis (Robinson and Wilson, 1994; Willacker et al., 2010; Hulsey et al., 2013). However, the generality of this pattern has largely been inferred from fishes that inhabit lentic, or lake-like, environments, such as sticklebacks, perch, arctic charr, and cichlids (Robinson and Wilson, 1994; Willacker et al., 2010; Hulsey et al., 2013; Skúlason et al., 1989; Schluter, 1993; Svanbäck and Eklöv, 2003; Meyer, 1990). Furthermore, most of these studies have examined microevolutionary processes of interspecific divergence. It remains unclear if benthic/pelagic divergence has commonly influenced macroevolutionary patterns within large clades of fishes inhabiting ecologically complex lotic, or riverine, systems.

Cyprinid fishes have radiated extensively within flowing water environments across eastern North America to exploit both benthic and pelagic habitats. Therefore, this group should provide an ideal study system to test whether the benthic to pelagic habitat

axis drives macroevolution in a species-rich group of fishes. Previous phylogenetic analyses have generated a general framework for understanding relationships among these fishes and have shown that most (>95%) of the cyprinid species inhabiting eastern North America form a strongly supported clade (Mayden, 1989; Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010; Hollingsworth and Hulsey, 2011). This clade is united by the osteological character of a small opening at the base of the skull known as the open posterior myodome (OPM). Previous studies have also generally agreed that a small clade of seven species, with two distributed in eastern North America (*Clinostomus elongatus* and *C. funduloides*) and five endemic to western North America (*Iotichthys phlegethontis*, *Mylocheilus caurinus*, *Pogonichthys macrolepidotus*, *Richardsonius balteatus*, and *R. egregius*) form the sister group to a much larger group of species that is primarily confined to eastern North America (Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010). Within this eastern radiation a strongly supported clade of around 200 predominantly pelagic species that display terminal mouths and feed generally from the mid-water is consistently recovered as arising following the initial divergence of several depauperate and strictly benthic lineages that display inferior mouths and often posses maxillary barbels (Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010). However, phylogenetic ambiguity remains in the branching order of these early benthic lineages, and the phylogenetic affinities of many species that are thought to lie within the predominately pelagic clade have not been resolved (Mayden, 1989; Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010; Hollingsworth and Hulsey, 2011). Therefore, a much more exhaustively sampled phylogeny combined with data on benthic/pelagic habitat use should facilitate a more robust phylogenetic examination of whether this ecological axis has influenced diversification within OPM cyprinids.

Habitat divergence clearly promotes coexistence in many lotic systems. For instance, OPM cyprinids often form complex communities consisting of up to 15 species that partition the water column into vertically stratified foraging zones (Page and Burr, 2011; Baker and Ross, 1981; Gorman, 1988a,b). Furthermore, small, insectivorous or omnivorous fishes from other groups are relatively rare in the pelagic zone of rivers and streams in eastern North America (Page and Burr, 2011). Therefore, the first transition from a benthic to pelagic habitat in OPM cyprinids likely represented the invasion of a sparsely occupied adaptive zone that could have resulted in a period of accelerated diversification (Simpson, 1953; Schluter, 2000; Losos, 2010; Hulsey and Hollingsworth, 2011). Given the apparent influence of the benthic/pelagic axis on community structure**,** mapping this habitat divergence onto the OPM phylogeny could highlight its role in generating species diversity.

Hypotheses addressing the ecological mechanisms that have influenced historical patterns of diversification can now be examined using robustly sampled molecular phylogenies and applying methods that examine phylogenetic tree shape (Glor, 2010). Acceleration in diversification rate is often thought to result from rapid divergence following invasion of open adaptive zones in groups ranging from vertebrates to prokaryotes (Harmon et al., 2008; Rabosky and Lovette, 2008; Dumont et al., 2012; Fordyce, 2010; Morlon et al., 2012). Yet within freshwater fishes, several tests for

ecologically associated bursts of diversification have failed to reject a constant rate of cladogenesis (Day et al., 2008; Hulsey et al., 2010; Day et al., 2013). However, OPM cyprinids could have experienced an exceptional period of lineage diversification following their initial transition from benthic to pelagic habitats.

In this study, we generated the most thoroughly sampled, species-level phylogenetic hypotheses for OPM cyprinids using DNA sequence data from both a mitochondrial and nuclear marker. We then used ancestral state reconstruction and fossil-calibrated divergence time estimates to infer the history of benthic/pelagic habitat use across our phylogenetic reconstructions. Using several independent methods, we addressed the question: Was the first major evolutionary shift from benthic to pelagic habitats in eastern North America followed by a period of accelerated lineage diversification in OPM cyprinids?

Methods

Phylogenetic and divergence time analyses

We used a combination of FishBase (2000) and Page and Burr (2011) to generate a list of the currently recognized species of OPM cyprinids. Using DNA sequences downloaded from GenBank combined with new sequence data, we constructed manually aligned matrices for the mitochondrial Cyt*b* and nuclear Rag1 loci. To obtain new sequence data, we first used DNAeasy Tissue Extraction Kits (Qiagen, Valencia, CA) to extract genomic DNA from tissue samples. Cyt*b* was amplified using primers from Schmidt and Gold (1993). Rag1 was amplified using primers from Lopez et al. (2004). DNA sequencing was performed at the University of Washington's High Throughput Genomics Unit using the PCR primers and an internal primer to sequence Rag1, IF4: 5'-TGAGAAGGCAGTGAGGTTTT-3'. We created contiguous sequence files from directional sequence reads using Sequencher 4.8 (Gene Codes, Ann Arbor, MI, USA) and coded heterozygous sites in the Rag1 alignment as ambiguous. The Cyt*b* alignment (1060 -1140 bp) included data for 223 of the 238 (94%) extant OPM taxa. The Rag1 alignment (1440 - 1518 bp) included data for 187 of the 238 (79%) extant OPM taxa. Our sampling includes taxa from throughout the geographic range of the clade with no obvious sampling bias between benthic and pelagic taxa and is given in the attachments (File 2.1). We deposited all new sequence data on GenBank [GenBank: KC763652-KC763776] (File 2.1). This includes 47 novel Cyt*b* sequences and 78 novel Rag1 sequences.

We estimated the phylogenies and divergence times for each of our two loci separately utilizing BEAST v1.7.1 (Drummond et al., 2012). We first defined codon positions in our two alignments using MacClade v4.07 (Maddison and Maddison, 2000) and then assigned the best model based on AIC scores calculated by jModelTest v0.1 (Posada, 2008) to each gene's codon sites in BEAUti v1.7.1 (Drummond et al., 2012). Substitution rate, rate heterogeneity, and base frequency parameters were treated as unlinked across partitions. We used a birth-death speciation prior for our tree models. The OPM clade containing *Mylocheilus caurinus*, *Pogonichthys macrolepidotus*, *Clinostomus funduloides*, *Clinostomus elongatus, Iotichthys phlegethontis*, *Richardsonius*

balteatus, and *Richardsonius egregious* was included as the outgroup to the eastern OPM radiation in our phylogenetic analyses (Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010). We conducted a single heuristic likelihood tree search on each gene alignment using RAxML v7.0.4 (Stamatakis et al., 2005) to generate starting trees for our MCMC runs. We also concatenated the two alignments and ran a phylogenetic analysis on this combined matrix using the same models specified in the individual gene analyses.

To estimate divergence times within the OPM radiation, we used an uncorrelated lognormal molecular clock model to temporally calibrate our two gene trees and concatenated phylogeny. Based on previous results (Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010), we constrained the monophyly of *Mylocheilus caurinus* and *Pogonichthys macrolepidotus* and then defined a fossil-calibrated prior distribution on the age of their MRCA. The fossil species *Mylocheilus whitei* was used to infer a minimum age estimate for this split. This fossil is a pharyngeal arch with a short anterior limb and thick internal ridges displaying small canals and pores, as well as molariform dentition, which are characters that are diagnostic for *Mylocheilus* (Smith and Cossel, 2002). The fossil was recovered from a geological layer representing the Clarendonian/Hemphillian boundary at approximately 9 million years ago (mya) (Smith and Cossel, 2002). We therefore specified a lognormal prior for the MRCA of *Mylocheilus caurinus* and *Pogonichthys macrolepidotus* with a mean and standard deviation of 1 mya and offset by 9 mya. The root node ages of our trees were constrained using a uniform prior with an upper bound of 75 mya based on recent MRCA age estimates of Cyprinidae (Near et al., 2012). Our MCMC chains were run for 2.0×10^7 generations, with trees and parameter estimates logged every 1.0×10^4 generations. We then ran each MCMC search five times using the CIPRES Science Gateway (Miller et al., 2010). The first 10% of each run was discarded as the burn-in. Subsequently, we examined ESS values in TRACER v1.5 (Rambaut and Drummond, 2007) over the remainder of the run to ensure convergence of parameter estimates. We combined log and tree files using LogCombiner v1.7.1 and Tree Annotator v1.7.1 (Drummond et al., 2012) to calculate the maximum clade credibility (MCC) tree for each locus and the concatenated analysis. All trees are deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S15034).

Habitat designations and ancestral state reconstruction

To estimate the most likely ancestral node representing the first benthic to pelagic shift of OPM cyprinids in eastern North America, we first designated extant taxa as benthic (0) or pelagic (1) based on a combination of morphological and ecological characteristics (File 2.1). Taxa coded as benthic display some combination of the following characteristics: 1) mouth is located ventrally 2) possess barbels 3) exhibit a spiraled gut 4) build benthic nests and 5) feed primarily on benthic food items. We coded taxa that do not display any of these traits as pelagic. We then used Pagel's (1994) single-rate Markov model of binary character evolution and assumed equal transition probabilities to reconstruct benthic/pelagic ancestral states using the package *ape* v3.0

(Paradis et al., 2004) in R (R development core team, 2013). We considered the most ancestral node inferred to have a greater than 50% probability of being pelagic to represent the initial benthic to pelagic transition in our phylogenies. This method of habitat coding and ancestral state reconstruction provided a conservative approach for inferring the phylogenetic placement of the benthic/pelagic shift on our gene trees. All taxa that could possibly be benthic were coded as such. Furthermore, all taxa diverging before the 'transition node' are unambiguously benthic. Therefore, any possible alternative codings would include more pelagic taxa in the focal clade and would result in the same nodes in our gene trees being recovered as the 'transition node'.

Diversification rate analyses

We employed two strategies to test for a period of accelerated diversification following the first benthic to pelagic transition in the OPM radiation of eastern North America. First, we used the entire Cyt*b* gene tree and conducted the relative cladogenesis (RC) test (Purvis et al., 1995) to identify significantly diverse subclades using the R package *geiger* v1.0 (Harmon et al., 2008). Using a homogeneous model of cladogenesis, this test examines the number of lineages alive just before a node and the number of lineages descending from the node and calculates the probability that the node has as many descendants as it has empirically (Purvis et al., 1995; Harmon et al., 2008). We also used the parametric rates comparison (PRC) test of Shah et al. (2013) implemented in the R package *iteRates* v3.0. This method iterates across all subtrees within the phylogeny that contain at least 6 edges, fits distributions to the vector of branch lengths within each subtree, and compares the likelihood that the vector of branch lengths from each subtree is best modeled as being drawn from the same, or different, distribution as the remainder of the tree. We fit an exponential distribution to our vector of branching times in the PRC analysis. Both of these methods, RC and PRC, assume complete taxon sampling [Purvis et al., 1995; Shah et al., 2013). Therefore, we only conducted these tests on the more robustly sampled Cyt*b* MCC gene tree. Both taxonomic inflation and using a single individual per species can bias these types of analyses because of obscured patterns of cladogenesis at the tips of the gene tree (Rabosky and Lovette, 2008; Isaac et al., 2004). Therefore, we truncated the most recent five million years from the Cyt*b* tree before conducting these tests using the *treeTrim* function in *iteRates* v3.0.

We next focused on the predominately pelagic clade subtending the 'transition node', which we refer to as the 'focal clade'. We used several independent analyses to test for an early period of rapid cladogenesis in this clade using a combination of the R packages *geiger* v1.0 (Harmon et al., 2008) and *laser* v2.2 (Rabosky, 2006). We first used Pybus and Harvey's (2000) constant rates test on our two focal clade phylogenies. This test is frequently called the γ test based on its test statistic, γ, which is distributed as a standard normal variable under a pure-birth process (Pybus and Harvey, 2000). Values of $\gamma < -1.645$ are considered significant deviations from pure-birth with diversification events clustered towards the base of a tree.

However, our phylogenies only included 91% (Cyt*b*) and 74% (Rag1) of the recognized species diversity in the focal clade and incomplete taxon sampling will bias

our calculations of γ because incomplete lineage sampling prunes tips from the tree, thereby inflating the branch lengths in the recent past (Pybus and Harvey, 2000). To correct for this bias, we employed Pybus and Harvey's (2000) Monte Carlo constant rates test (MCCR test), where the critical value for rejecting a constant rate (at $\alpha = 0.05$) is calculated by examining the distribution of γ for simulated trees that include incomplete taxon sampling. Our null distribution of γ was calculated from 1 million simulated pure-birth trees of 192 taxa, or the number of described species that belong to the genera comprising the focal clade. Our simulated trees were corrected for the number of taxa missing in our reconstructed Cyt*b* and Rag1 phylogenies by randomly pruning 18 and 50 taxa from each simulated tree, respectively. All phylogenetic simulations used a modification of the *birthdeath.tree* function in *geiger* v1.0 to ensure that the trees had the desired statistical properties [see (Fordyce, 2010) for details]. Additionally, we calculated the 'tree deviation' statistic (Fordyce, 2010), which can have greater power to detect accelerated diversification early in the history of a tree by examining if lineages have accumulated at a greater rate than predicted by a null distribution. The null distribution for the 'tree deviation' was calculated from 1 million simulated pure-birth trees with incomplete taxon sampling. We also generated a lineage through time plot (LTT) for our focal clade based on the more thoroughly sampled Cyt*b* MCC gene tree to compare to a distribution of 10,000 simulated pure-birth LTT plots.

We then used a likelihood-based approach to test for a deviation from a constant-rate pattern of diversification in the focal clade. We fit two constant-rate models (pure-birth, birth-death) and three variable-rate models (density dependent logistic, density dependent exponential, and Yule 2-rate) to the vector of branching times from the two focal clade phylogenies. To determine the best-fit model for our data and to account for incomplete taxon sampling, we used the method proposed by Rabosky [56]. This method compares the observed ΔAIC between the best-fit constant-rate model and the best-fit variable-rate model of a focal tree to the 0.95 quantile of a null distribution of ΔAIC values calculated from 1 million simulated pure-birth phylogenies with incomplete lineage sampling. The constant rates test (Pybus and Harvey, 2000), tree deviation (Fordyce, 2010), and model fitting approach of Rabosky (2006) were applied to the MCC gene trees for Cyt*b* and Rag1, and also across all post burn-in trees from the BEAST analysis.

Results

Phylogenetic reconstruction

Phylogenetic analysis of the cytochrome *b* (Cyt*b*) and recombination activating gene 1 exon 3 (Rag1) loci provided substantial resolution of relationships among members of the OPM radiation. These maximum clade credibility trees (MCC) from these analyses are given in the attachments (File 2.2). Both MCC gene trees and the MCC concatenated analysis included moderately to well-supported clades (>90% posterior probability (pp)) containing the benthic genera *Campostoma*, *Exoglossum*, *Nocomis*, and *Rhinichthys* as the earliest diverging OPM lineages in eastern North America (Figure 2.1

and File 2.2). Both gene trees recovered the remaining eastern benthic genera *Dionda*, *Erimystax*, *Macrhybopsis*, *Phenacobius*, and *Platygobio* as diverging before the diversification of the strongly supported (100% pp) focal clade (see Methods) that is dominated by pelagic species and accounts for approximately 80% of extant OPM diversity. This general topology was strongly supported in the concatenated analysis as well (Figure 2.1). Some phylogenetic relationships within the predominately pelagic focal clade were variable between our two gene trees and the concatenated analysis, with many clades receiving varying levels of posterior support. For instance, we consistently recovered poorly resolved nodes and short internode branch lengths at the base of the focal clade.

Ancestral state reconstruction and divergence time estimates

Ancestral state reconstruction points to a similar node in both of our gene trees as representing the initial shift from benthic to pelagic habitat utilization in the eastern OPM radiation (Figure 2.2 and File 2.2). In the more thoroughly sampled Cyt*b* gene tree, this shift is inferred to have occurred along a branch leading to the most recent common ancestor (MRCA) of a strongly supported focal clade. We considered the node representing this MRCA as the transition node and conducted our tests for variation in diversification rate and an excess of early lineages after the habitat shift with respect to this node. In the Rag1 gene tree, we recovered that the shift occurred at a slightly more ancestral node. This node was subtended by a clade containing the same set of genera as the Cyt*b* focal clade plus its sister group *Macrhybopsis* spp. + *Platygobio gracilis*. However, this node received poor support (62% pp) in the Rag1 gene tree. Furthermore, the subsequent node in the Rag1 tree that was subtended by the same set of genera as the Cyt*b* focal clade was strongly supported (100% pp) and had a higher likelihood of being pelagic. Preliminary analyses suggested that using this node as opposed to its poorly supported ancestral node had little impact on our calculation of diversification rate and tree shape statistics. Therefore, we considered this to be the transition node for the Rag1 topology as well. We estimated the age of the transition nodes to be 33 mya (95% highest posterior density: 16-47 mya) based on Cyt*b* (Figure 2.3), 27 mya (95% highest posterior density: 16-61 mya) based on Rag1.

Following the initial benthic to pelagic shift early in the history of the OPM radiation, ancestral state reconstruction recovered several instances of the re-evolution of benthicity within the pelagic focal clade (Figure 2.2). Examples of lineages that have re-evolved benthic specialization from within this pelagic clade include the barbeled genus *Hybopsis* and a sister species pair of barbeled *Cyprinella*, *C. labrosa* and *C. zanema.* Based on the more thoroughly sampled Cyt*b* MCC topology we recovered approximately 6 transitions back to benthic habitat use during the history of the clade. After these transitions back to benthicity, transitions back to pelagic habitats were very rare (Figure 2.2).

Diversification rate analyses

Our strategies used to examine topological imbalance and variation in diversification rate across the entire OPM phylogeny supported the hypothesis that there was accelerated diversification following the initial benthic to pelagic transition in the OPM radiation. The relative cladogenesis (RC) test identified 15 nodes associated with significantly diverse subclades in our Cyt*b* MCC gene tree (Figure 2.2A). These nodes included two that are immediately ancestral to the transition node at the base of the focal clade, the transition node itself, and 12 consecutive descendent nodes. The parametric rates comparison (PRC) analysis marginally supported a model with a higher diversification rate in the predominately pelagic focal clade relative to the remainder of the Cyt*b* gene tree ($p = 0.06$) (Figure 2.3A). We also found that three nodes immediately following the transition node represented clades that were significantly more likely to be modeled as having a greater diversification rate relative to the remainder of the tree at $\alpha = 0.1$.

Likewise, our examinations of deviations from a constant rate of cladogenesis as compared to randomly generated pure-birth topologies also supported the hypothesis that there was a burst of diversification coincident with the initial shift to a pelagic habit. The γ statistic was significant for our focal clade in both the Cyt*b* and Rag1 MCC gene trees (Table 2.1) indicating an excess of early lineages in this clade. The Cyt*b* LTT plot for the focal clade lay largely outside the 95% confidence intervals for nearly the entire history of 10,000 simulated pure-birth trees (Figure 2.3B). The Monte Carlo constant rates (MCCR) analyses indicated a strong deviation from a pure-birth process based on Cyt*b* and the MCCR analyses marginally supported this deviation in the Rag1 gene tree (Table 2.1). Tree deviation scores were also significant on our two gene trees, again indicating an excess of early lineages in our focal clade (Table 2.1). Finally, variable-rate models provided a significantly better fit than constant-rate models to the observed vectors of branching times within the predominately pelagic focal clade (Table 2.1). We obtained similar results when we applied these test to the 9005 post burn-in trees (Figure 2.4).

Discussion

The initial evolutionary transition from benthic to pelagic habitats by OPM cyprinid fishes likely had a significant impact on the diversification of this hyper-diverse clade of fishes. Our two phylogenetic hypotheses, coupled with ancestral state reconstructions and divergence time estimates, indicated that benthic forms dominated the early history of the eastern OPM radiation. This group then gave rise to a predominately pelagic clade that began diversifying around 30 mya and contains ~80% of extant OPM species. Our tests for increased diversification rate all highlighted the particular region in the phylogeny where the initial benthic to pelagic habitat shift is inferred to have occurred. With our thoroughly sampled phylogenies, we were also able to reject a pattern of constant-rate cladogenesis in favor of models that are consistent with a period of accelerated diversification after this habitat shift. As such, this initial benthic

to pelagic transition by OPM cyprinids likely represented evolution into an open adaptive zone that resulted in a period of rapid lineage accumulation (Simpson, 1953; Schluter, 2000; Losos, 2010).

In the area of the phylogeny immediately following this inferred benthic to pelagic transition, we were not able to confidently resolve relationships among lineages. This region has often been unresolved in other studies of OPM evolution (Mayden, 1989; Simons et al., 2003; Bufalino and Mayden, 2010; Houston et al., 2010; Hollingsworth and Hulsey, 2011). The explosive diversification in this region of the tree likely has contributed to this phylogenetic ambiguity (Hollingsworth and Hulsey, 2011). Future phylogenetic studies based on large, multi-locus datasets that utilize a species tree framework could potentially help to resolve these problematic areas of the OPM phylogeny (Hollingsworth and Hulsey, 2011). However, determining the exact branching order of lineages whose divergence is coincident with major ecological shifts and periods of rapid diversification might be generally difficult.

The major shift from benthic to pelagic habitats in OPM cyprinids should not be considered in isolation from the other freshwater fish diversity in eastern North America. The diversification in the predominately pelagic focal clade that began around 30 mya coincides with the estimated age of the darter (Percidae: Etheostomatinae) radiation (Near et al., 2011). Darters are another endemic North American freshwater group of around 250 benthic fishes that often co-occur with OPM species (Page and Burr, 2011; Near et al., 2011). Cyprinids and darters together dominate the abundance and species diversity in most eastern North America fish assemblages (Page and Burr, 2011). With the exception of a few omnivorous species, OPM cyprinids and darters are both primarily insectivorous and compete for similarly sized prey (Knight et al., 2008). Therefore, darter diversification might have further reduced eco-evolutionary opportunities within benthic habitats and influenced the shift of the OPM lineage into the relatively competitor-free pelagic zone. A macroevolutionary interaction between these two lineages could have contributed to the observed pelagic burst of OPM diversity.

Interspecific competition, however, might not have been the only mechanism driving the rapid diversification of pelagic OPM cyprinids. For instance, there is also an increase in the presence of male nuptial coloration and sexual dichromatism in the more visually dependent pelagic OPM species relative to their benthic relatives that rely more extensively on chemical cues during foraging (Page and Burr, 2011). Given this, an increase in visually mediated sexual selection could have also played a role in the diversification of the pelagic OPM cyprinids (Barraclough et al., 1995; Kazancıoğlu et al., 2009). Additionally, an interaction between ecological opportunity and sexual selection might have driven the increased rate of lineage accumulation that followed the first benthic to pelagic transition in this group (Wagner et al., 2012).

Evolution along the benthic/pelagic habitat axis appears to have played a critical role in generating the impressive species numbers of OPM cyprinids inhabiting the lotic systems of eastern North America. Future studies of other freshwater fish groups that combine ecological data with more thoroughly sampled phylogenies and examinations of temporal shifts in diversification could provide additional evidence that divergence along this axis has repeatedly influenced fish macroevolution. Our results indicate that the

influence of this habitat axis is clearly not restricted to lentic environments (Robinson and Wilson, 1994; Willacker et al., 2010; Hulsey et al., 2013; Skúlason et al., 1989; Schluter, 1993; Svanbäck and Eklöv, 2003; Meyer, 1990). Instead, the benthic/pelagic axis of diversification appears to be a ubiquitous generator of biodiversity across disparate freshwater ecosystems.

References

Baker, J.A., Ross, S.T., 1981. Spatial and temporal resource utilization by southeastern cyprinids**.** *Copeia* 1981, 178-189.

Barbour, C.D., Miller, R.R., 1978. A revision of the Mexican cyprinid fish genus *Algansea*. *Misc. Pub. Mus. Zool. Univ. Mich.* 155, 1-72.

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

Bufalino, A.P., Mayden, R.L., 2010. Phylogenetic evaluation of North American Leuciscidae (Actinopterygii: Cypriniformes: Cyprinoidea) as inferred from analyses of mitochondrial and nuclear DNA sequences. *Syst. Biodivers.* 8, 493-505.

Chernoff, B., Miller, R.R., 1986. Fishes of the *Notropis calientis* complex with a key to the southern shiners of México. *Copeia* 1986, 170-183.

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

Day, J.J., Peart, C.R., Brown, K.J., Friel, J.P., Bills, R., Moritz, T., 2013. Continental diversification of an African catfish radiation (Mochokidae: *Synodontis*). *Syst. Biol.* 62, 351-365.

Domínguez-Domínguez, O., Pompa-Domínguez, A., Doadrio, I., 2007. A new species of the genus *Yuriria* Jordan & Evermann, 1896 (Actinopterygii, Cyprinidae) from the Ameca Basin of the Central Mexican Plateau. *Graellsia* 63, 259-271.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969-1973.

Dumont, E.R., Dávalos, L.M., Goldberg, A., Santana, S.E., Rex, K., Voigt, C.C., 2012. Morphological innovation, diversification, and invasion of a new adaptive zone**.** *Proc. R. Soc. B* 279, 1797-1805.

Etnier, D.A., Starnes, W.C., 1993. The fishes of Tennessee. University of Tennessee Press, Knoxville, TN.

Fordyce, J.A., 2010. Host shifts and evolutionary radiations of butterflies. *Proc. R. Soc. B* 277, 3735-3743.

Fordyce, J.A., 2010. Interpreting the gamma statistic in phylogenetic diversification rate studies: a rate decrease does not necessarily indicate an early burst. *PLoS One* 5, e11781.

Froese, R., Pauly, D., 2000. FishBase 2000: concepts, design and data sources. ICLARM, Los Baños.

Glor, R.E., 2010. Phylogenetic insights on adaptive radiation. *Annu. Rev. Ecol. Evol. Syst.* 41, 251-270.

Gorman, O.T., 1988. The dynamics of habitat use in a guild of Ozark minnows. *Ecol. Monogr.* 58, 1-18.

Gorman, O.T., 1988. An experimental study of habitat use in an assemblage of Ozark minnows. *Ecology* 69, 1239-1250.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., Challenger, W., 2008. GEIGER: investigation evolutionary radiations. *Bioinformatics* 24, 129-131.

Harmon, L.J., Melville, J., Larson, A., Losos, J.B., 2008. The role of geography and ecological opportunity in the diversification of day geckos (*Phelsuma*). *Syst. Biol.* 57, 562-573.

Hollingsworth Jr., P.R., Hulsey, C.D., 2011. Reconciling gene trees of eastern North American minnows. *Mol. Phylogenet. Evol.* 61, 149-156.

Houston, D.D., Shiozawa, D.K., Riddle, B.R., 2010. Phylogenetic relationships of the western North American cyprinid genus *Richardsonius*, with an overview of phylogeographic structure. *Mol. Phylogenet. Evol.* 55, 259-273.

Hulsey, C.D., Hollingsworth Jr., P.R., 2011. Do constructional constraints influence cyprinid (Cyprinidae: Leuciscinae) craniofacial coevolution? *Biol. J. Linn. Soc.* 103, 136-146.

Hulsey, C.D., Hollingsworth Jr., P.R., Fordyce, J.A., 2010. Temporal diversification of Central American cichlids. *BMC Evol. Biol.* 10, 279.

Hulsey, C.D., Roberts, R.J., Loh, Y.H.E., Rupp, M.F., Streelman, J.T., 2013. Lake Malawi cichlid evolution along a benthic/limnetic axis. *Ecol. Evol.* 3, 2262-2272.

Isaac, N.J.B., Mallet, J., Mace, G.M., 2004. Taxonomic inflation: its influence on macroecology and conservation. *Trends Ecol. Evol.* 19, 464-469.

Jenkins, R.E., Burkhead, N.M., 1994. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, MD.

Kazancıoğlu, E., Near, T.J., Hanel, R., Wainwright, P.C., 2009. Influence of sexual selection and feeding functional morphology on diversification rate of parrotfishes (Scaridae). *Proc. R. Soc. B* 276, 3439-3446.

Knight, R.R., Gregory, M.B., Wales, A.K., 2008. Relating streamflow characteristics to specialized insectivores in the Tennessee River Valley: a regional approach. *Ecohydrology* 1, 394-407.

Lopez, J.A., Chen, W.J., Orti, G., 2004. Esociform phylogeny. *Copeia* 2004, 449-464.

Losos, J.B., 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* 175, 623-639.

Maddison, D.R., Maddison, W.P., 2000. Macclade. Sinauer Associates, Inc., Sunderland, MA.

Mayden, R.L., 1989. Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: Cypriniformes). *Misc. Pub. Univ. Kans. Mus. Nat. Hist.* 80, 1-189.

Meyer, A., 1990. Ecological and evolutionary consequences of the trophic polymorphism in *Cichlasoma citrinellum* (Pisces: Cichlidae). *Biol. J. Linn. Soc.* 39, 279-299.

Miller, R.R., Minckley, W.L., Norris, S.M., 2005. Freshwater fishes of Mexico. University of Chicago Press, Chicago, IL.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop. New Orleans, 1–8.

Morlon, H., Kemps, B.D., Plotkin, J.B., Brisson, D., 2012. Explosive radiation of a bacterial species group. *Evolution* 66, 2577-2586.

Near, T.J., Bossu, C.M., Bradburd, G.S., Carlson, R.L., Harrington, R.C., Hollingsworth Jr., P.R., Keck, B.P., Etnier, D.A., 2011. Phylogeny and temporal diversification of darters (Percidae: Etheostomatinae). *Syst. Biol.* 60, 565-595.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M., Smith, W.L., 2012. Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl. Acad. Sci. USA* 109, 13698-13703.

Page, L.M., Burr, B.M., 2011. Peterson field guide to freshwater fishes, second edition. Houghton Mifflin, Boston, MA.

Pagel, M., 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. B* 255, 37–45.

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

Pflieger, W.L., 1997. The fishes of Missouri. Missouri Department of Conservation, Jefferson City, MO.

Posada, D., 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253-1256.

Purvis, A., Nee, S., Harvey, P.H., 1995. Macroevolutionary influence from primate phylogeny. *Proc. R. Soc. B* 260, 329-333.

Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proc. R. Soc. B* 267, 2267-2272.

R Development Core Team, 2013. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. [http://www.R-project.org/]

Rabosky, D.L., 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol. Bioinform. Online* 2, 257-260.

Rabosky, D.L., 2006. Likelihood methods for inferring temporal shifts in diversification rates. *Evolution* 60, 1152-1164.

Rabosky, D.L., Lovette, I.J., 2008. Density-dependent diversification in North American wood warblers. *Proc. R. Soc. B* 275, 2363-2371.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. [http://beast.bio.ed.ac.uk/Tracer]

Robinson, B.W., Wilson, D.S., 1994. Character release and displacement in fishes: a neglected literature. *Am. Nat.* 144, 596-627.

Rohde, F.C., Arndt, R.G., Foltz, J.W., Quattro, J.M., 2009. Freshwater Fishes of South Carolina. University of South Carolina Press, Columbia, SC.

Ross, S.T., 2001. Inland fishes of Mississippi. University Press of Mississippi, Jackson, MS.

Schluter, D., 1993. Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology* 74, 699–709.

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

Schmidt, T.R., Gold, J.R., 1993. Complete sequence of the mitochondrial cytochrome-b gene in the cherryfin shiner, *Lythrurus roseipinnis* (Teleostei, Cyprinidae). *Copeia* 1993, 880-883.

Schönhuth, S., Doadrio, I., Domínguez-Domínguez, O., Hillis, D.M., Mayden, R.L., 2008. Molecular evolution of southern North American Cyprinidae (Actinopterygii), with the description of the new genus *Tampichthys* from central México. *Mol. Phylogenet. Evol.* 47, 729-756.

Shah, P., Fitzpatrick, B.M., Fordyce, J.A., 2013. A parametric method for assessing diversification-rate variation in phylogenetic trees. *Evolution* 67, 368-377.

Simons, A.M., Berendzen, P.B., Mayden, R.L., 2003. Molecular systematics of North American phoxinin genera (Actinopterygii: Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. *Zool. J. Linn. Soc.* 139, 63-80.

Simpson, G.G., 1953. The major features of evolution. Columbia University Press, New York, NY.

Skúlason, S., Noakes, D.L.G., Snorrason, S.S., 1989. Ontogeny of trophic morphology in four sympatric morphs of arctic charr *Salvelinus alpinus* in Thingvallavatn, Iceland. *Biol. J. Linn. Soc.* 38, 281–301.

Smith, G.R., Cossel Jr., J., 2002. Fishes from the late Miocene Poison Creek and Chalk Hills formations, Owyhee County, Idaho. In: Akersten, W.A., Thompson, M.E., Meldrum, D.J., Raup, R.A., McDonald, H.G. (Eds.), And whereas...papers on the vertebrate paleontology of Idaho honoring John A. White, Volume 2. Idaho Museum of Natural History, Pocatello, ID, pp. 23-35.

Stamatakis, A., Ludwig, T., Meier, H., 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* 21, 456–463.

Svanbäck, R., Eklöv, P., 2003. Morphology dependent foraging efficiency in perch: a trade-off for ecological specialization? *Oikos* 102, 273-284.

Wagner, C.E., Harmon, L.J., Seehausen, O., 2012. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487, 366-369.

Willacker, J.J., von Hippel, F.A., Wilton, P.R., Walton, K.M., 2010. Classification of threespine stickleback along the benthic-limnetic axis. *Biol. J. Linn. Soc.* 101, 595-608.

CHAPTER III

ADAPTIVE RADIATION ALONG A BENTHIC/PELAGIC AXIS IN NORTH AMERICA'S MOST DIVERSE, ENDEMIC CLADE OF FRESHWATER FISHES

Abstract

The theory of adaptive radiation predicts simultaneous bursts of cladogenesis and phenotypic evolution driven by differential exploitation of a set of resources. Previous studies have rarely recovered evidence for both. Furthermore, these studies generally fail to test for an association between phenotypic variation and resource partitioning. Building on the results of previous work that suggest a period of rapid diversification followed a major benthic to pelagic transition in open posterior myodome (OPM) cyprinid fishes, we use a multi-faceted approach to further test for evidence of an adaptive radiation along this ecological axis in this diverse clade. We recover a strong relationship between variation in an eco-morphological trait, jaw protrusion angle (JPA), and vertically segregated foraging zones within a complex OPM assemblage. We then identify a number of individual morphological traits comprising the cyprinid jaw apparatus that are highly correlated with JPA. Model fitting analyses of morphological rate variation suggest that a number of these traits experienced periods of rapid diversification following the benthic to pelagic shift in the OPM clade. Therefore, this ecological axis likely provided the setting for an adaptive radiation of freshwater fishes that diversified in trophic morphology to variably exploit habitat space throughout the water column.

Introduction

Ecological opportunity can fuel adaptive radiations and result in simultaneous periods of accelerated lineage diversification and explosive phenotypic evolution (Simpson, 1944, 1953; Schluter, 2000; Yoder et al., 2010; Losos, 2010). One potential source of ecological opportunity is entrance into an open adaptive zone (Simpson, 1944, 1953). There are a number of studies that suggest that phylogenetic bursts of diversification have followed the evolution of a clade into a novel adaptive zone (e.g., Harmon et al., 2008a; Fordyce, 2010). However, despite the apparent prevalence of early periods of exceptional lineage diversification across disparate clades, there appear to be few examples of coincident burst of phenotypic evolution that are expected under the model of adaptive radiation (Simpson, 1944, 1953; Harmon et al., 2010). Furthermore, previous studies have generally failed to explicitly link phenotypic traits with resource partitioning and to provide convincing evidence for periods of rapid evolution in those traits that are coincident with rapid lineage diversification (Dumont et al., 2012).

An earlier study on the evolutionary ecology of the North American endemic, open posterior myodome (OPM) cyprinid fishes presented evidence for a burst in lineage diversification after a major shift from benthic to pelagic habitats early in the history of this clade (Hollingsworth et al., 2013). Although divergence along a benthic/pelagic axis seems to be ubiquitous in lentic fishes (Robinson and Wilson, 1994; Schluter, 2000; Willacker et al., 2010; Hulsey et al., 2013), it remains to be tested whether this habitat gradient could drive adaptive radiations in diverse clades of lotic, freshwater fishes. For

instance, for a lineage of historically benthic fishes, the water column could serve as a novel, third dimension of habitat space that is capable of being further subdivided into vertically segregated foraging zones (DeVries, 1988; DeVries et al., 2012). However, it is not clear what phenotypic variation might allow for this resource partitioning and, therefore, be expected to have experienced a burst of evolution following the pelagic shift if this ecological transition triggered an adaptive radiation of OPM cyprinids.

One possible trait is jaw protrusion angle (JPA). Teleost fishes have evolved the ability to protrude their oral jaws during feeding at least three independent times, including within Cypriniformes (Westneat, 2004). Jaw protrusion, in turn, has been suggested to be a key innovation facilitating the ecological and phylogenetic diversification of disparate fish lineages (Schaffer and Rosen, 1961; Lauder, 1982; Ferry-Graham and Lauder, 2001; Konstantinidis and Harris, 2011). A commonly held, though largely untested, assumption regarding fish ecomorphology is that JPA influences a fish's ability to exploit food items in different parts of the water column (Alexander, 1966, 1967; Gatz, 1979). However, quantification of JPA and comparative analyses of teleost JPA evolution have received surprisingly little attention in the literature (Cochran-Biederman and Winemiller, 2010). Therefore, the morphological, mechanical, and behavioral factors influencing the evolution of this complex ecomorphological trait and its relationship to habitat partitioning in diverse fish assemblages remain uncertain.

This is particularly true for cypriniform fishes that display a unique mechanism of jaw protrusion that is quite divergent from that of well-studied perciform fishes in two main ways (Alexander, 1966, 1967; Motta, 1984; Hulsey et al., 2010). First, jaw protrusion is driven in Cypriniformes through the mechanics of a novel, cranial ossification known as the kinethmoid (Hernandez et al., 2007). The kinethmoid attaches via ligaments to the palantine, maxilla, premaxilla, and neurocranium, and rotates 90-180 degrees dorsoventrally during protrusion of the premaxilla (Alexander, 1966; Ballintijn et al., 1972; Hernandez et al., 2007; Staab et al., 2012). Secondly, in perciforms all three major divisions of the adductor mandibulae (AM) muscle complex function only to close the jaws or during respiration (Osse, 1969). In cypriniforms, conversely, the AM complex is directly involved in jaw protrusion as the adduction of the maxilla by AM1 enhances the dorsoventral flipping of the kinethmoid that results in ventrally-directed protrusion of the premaxilla (Ballintijn et al., 1972; Hernandez et al., 2007). However, the muscle activity patterns operating during cypriniform jaw protrusion are complex and debated (Alexander, 1966; Ballintijn et al., 1972; Motta, 1984), and their effect on JPA is also unclear.

Despite the assertion by some authors that this mechanism of jaw protrusion is constrained towards higher values of JPA (Ballintjin et al., 1972), or more ventrally-directed jaw protrusion, OPM species are characterized by a large degree of interspecific variation in this trait. Species range from those that have inferior mouths that protrude towards the substrate and suggest benthic feeding to species with more terminal jaws that might be expected to feed higher in the water column (Page and Burr, 2011). Furthermore, previous studies on community assembly provide evidence that interspecific partitioning of vertically stratified foraging zones is a possible mechanism that promotes cyprinid co-occurrence in exceptionally diverse freshwater fish

assemblages across eastern North America (Mendelson, 1975; Baker and Ross, 1981; Gorman, 1988a,b). However, it has never been directly tested whether JPA and preferred foraging height in the water column are evolutionarily correlated. In the first part of this study, we perform the first explicit test of this relationship among 15 OPM species that inhabit the Little River, Blount Co., TN. If variation in JPA has evolved to more efficiently exploit vertical subdivisions of the water column, then recovering evidence of a burst of JPA diversification after the pelagic shift would suggest this trait played a key role in allowing this clade to adaptively radiate along this three-dimensional, ecological axis.

In the second part of this study we attempt to identify individual muscoskeletal components of the cyprinid oral jaw apparatus that are evolutionarily correlated with JPA to serve as proxies for JPA in further analyses. We then use a model fitting approach to evaluate the support for a period of exceptionally rapid diversification in these correlates of JPA immediately following the pelagic shift. This approach has two advantages. First, given the relative ambiguity regarding how the mechanism of cypriniform jaw protrusion operates (Alexander, 1966; Ballintijn et al., 1972; Motta, 1984), this could serve to identify characters whose evolution is involved in changing the angle of this protrusion. Secondly, these muscoskeletal measurements can be obtained from cleared and stained specimens, allowing for increased taxon sampling over JPA measurements that have to be obtained from live specimens. We identified six traits that might be expected to show significant correlations with JPA based on our limited knowledge of the mechanistic underpinnings of cypriniform jaw protrusion. These characters are AM1 mass, AM2 mass, AM3 mass, lower jaw length, ascending process (AP) of the premaxilla length, and kinethmoid length (Ballintijn et al., 1972; Hernandez et al., 2007; Hulsey et al., 2010; Staab et al., 2012).

To address the hypothesis that OPM cyprinids adaptively radiated after invading the previously unoccupied pelagic zone by vertically partitioning that new habitat space, we conducted three complimentary analyses. We first examined the association between JPA and preferred foraging height in the water column among a diverse community of OPM cyprinids. We then tested for a significant evolutionary correlation between a number of individual muscoskeletal traits and JPA in attempt to identify characters whose evolution explains variation in JPA. Finally, we assess the fit of alternative models of the rates of morphological evolution in these proxies for JPA before and after the major benthic to pelagic transition in the OPM clade (Hollingsworth et al., 2013). This study represents one of the most robust investigations of an adaptive radiation to be conducted (Dumont et al., 2012) and the results should elucidate the influence of the frequently cited benthic/pelagic axis of diversification on the macroevolution of a speciose clade of lotic, freshwater fishes.

Methods

Foraging height in the water column and JPA

In order to test for a significant evolutionary relationship between JPA and preferred foraging height in the water column, we quantified these two variables for 15 OPM species that co-occur in the Little River in Blount Co., TN. Snorkeling surveys were conducted to quantify habitat use at 6 sites along the Little River and we sampled each site 1-6 times. These surveys consisted of first identifying a riffle, run, pool complex approximately 25 m long and 5-10 m wide to serve as the sampling area. The sampling area was approached from the downstream side and transects were snorkeled back and forth across the area, advancing upstream, and identifying individual fishes for quantification of foraging height. Once an individual was identified to species, it was observed until it moved to inspect or ingest a potential food item from the benthos or drifting downstream. Its position in the water column was then recorded as an integer (1, 2, 3, or 4) on a dive slate. Individuals foraging directly from or within approximately 5 cm of the bottom were given a score of 1. The rest of the water column was visually divided into thirds. Individuals foraging in the bottom third of the water column were scored as 2, the middle third as 3, and the top third to the surface as 4. We only scored adult individuals that could be unambiguously identified to species. We attempted to avoid auto-correlation among schooling individuals by randomly identifying and scoring one individual per school. Snorkeling sessions lasted from approximately 30 minutes to 1 hour. We scored foraging height for 5-50 individuals per species (Table 3.1).

We quantified JPA for the same 15 species for which we gathered foraging height data. Individual fishes were captured using a standard seine net. An individual was then anesthetized using the common fish anesthetic MS-222 and positioned on a piece of waterproof graph paper with the eye centered on an intersection of grid lines and the caudal fin centered on the same horizontal line (Figure 3.1A). Next, pulling down the mandible with a pair of jeweler's forceps resulted in protrusion of the premaxilla and a mark was made using a 0.5 mm mechanical pencil at the tip of the maximally protruded premaxilla (Figures 3.1A,B). A line was then drawn using a straight edge from the grid intersection centering the eye to this mark. The angle between this line and the horizontal was measured as JPA (Figure 3.1B). We measured JPA on 7-30 individuals per species (Table 3.1).

We employed phylogenetically independent contrasts (PICs) (Felsenstein, 1985), calculated using the package *ape* v.3.0 (Paradis et al., 2004) in R (R Development Core Team, 2013), in order to test for a phylogenetically corrected relationship between JPA and foraging height across this community. We utilized the maximum clade credibility (MCC) cytochrome *b* (Cyt*b*) and recombination activating gene 1 exon 3 (Rag1) phylogenies from Hollingsworth et al. (2013) and pruned these trees down to these 15 Little River OPM species in order to calculate PICs. We then tested for a correlation that was forced through the origin between the PICs of mean JPA and median foraging height score on each of our two MCC phylogenies (Garland et al., 1992). We also tested this

correlation over 900 subsampled trees from the posterior distributions of Cyt*b* and Rag1 trees resulting from the Bayesian phylogenetic analyses in Hollingsworth et al. (2013).

Identifying functional morphological correlates of JPA

We next tested for a relationship between six morphological characters comprising the cyprinid jaw apparatus and JPA in order to identify functional, morphological proxies explaining evolutionary variation in JPA. We expanded our JPA dataset to include data for 40 OPM taxa. Sample sizes for this expanded JPA dataset consisted of 3-30 individuals per species. Specimens from the University of Tennessee Etnier Ichthyology Collection (UTEIC) were then utilized to measure the masses of the three individual AM muscles and the lengths of three osteological characters (lower jaw, AP, and kinethmoid) for the same 40 species included in the expanded JPA dataset. We dissected AMs and measured osteological characters on three individuals per species. The three individual AM muscles were dissected from an individual and placed in 1.5 mL tubes in 70% ethanol (Figure 3.1C). Before weighing each muscle it was removed from the tube and patted on a paper towel two times on each side to remove excess ethanol. The mass of each muscle was measured to the nearest 0.01 mg using a digital balance.

After dissection of the AM muscles, we cleared and stained the three specimens per species for quantification of osteological characters (Dingerkus and Uhler, 1977). Individual specimens were first worked from being saturated with 70% ethanol to 100% water. We then soaked individual specimens in a Trypsin bath until partially digested and stained them with a solution of 0.1 mg of Alizarin red stain per 100 mL potassium hydroxide. Specimens were placed in 1% potassium hydroxide for final clearing. Cleared and stained specimens are accessioned into the UTEIC.

We measured the lengths of our three osteological characters to the nearest 0.1 mm using an Olympus dissecting scope with an ocular micrometer (Figure 3.1D). Lower jaw length was measured from the quadrate/articular joint to the tip of the mandible. We then protruded an individual's premaxilla with a pair of forceps and measured the length of the AP from a lateral view and the length of the kinethmoid from a dorsal view. The standard length of each specimen was also obtained to use for size correction.

We applied the following transformations to our morphological data prior to conducting comparative analyses. As mass scales with the third power of length, we calculated the cube root of each of the three AM masses. The cube rooted AM masses, 3 osteological measurements, and standard lengths were then $\log_{10}$-transformed. Using these transformed values, we size corrected the six morphological traits by standard length using the *phyl.resid* function in the *phytools* v.0.3-72 package (Revell, 2009; Revell, 2012) in R along with the Cyt*b* and Rag1 MCC phylogenies from Hollingsworth et al. (2013) pruned down to include the 40 species being analyzed. JPA was untransformed and not size-corrected. We utilized PICs to test for coevolution between JPA and each these six characters. PICs were calculated using *ape* v.3.0 (Paradis et al., 2004) and the pruned Cyt*b* and Rag1 MCC phylogenies that were used during size-correction. We then tested for a correlation that was forced through the origin between the various sets of PICs (Garland et al., 1992). We also conducted these analyses over

900 subsampled Cyt*b* and Rag1 trees from the posterior distributions of Bayesian searches in Hollingsworth et al. (2013). The size-corrections using *phyl.resid* (Revell, 2009; Revell, 2012) were conducted on each of the individual 900 trees prior to calculating PICs for that tree.

Modeling rates of morphological evolution in correlates of JPA

In order to assess the support for a burst of morphological evolution in characters correlated with JPA after the OPM invasion of the pelagic zone, we first expanded our dataset to include the three AM masses and three osteological measurements for 141 OPM taxa. This included 23 species diverging before, and 118 diverging after, the benthic to pelagic transition (Hollingsworth et al., 2013). We pruned the Cyt*b* MCC and post-burn in phylogenies from Hollingsworth et al. (2013) to match the species in this expanded morphological dataset. Our expanded dataset for Rag1 included 16 less species in the clade arising after the pelagic shift. Morphological measurements were then transformed and size corrected as in the previous section of the Methods (see above).

If the OPM clade adaptively radiated by evolving variation in JPA that allowed them to partition the water column then we would expect to see both a faster rate of evolution after the pelagic shift in our correlates of JPA, as well as, a slow down in these rates through time after the shift as pelagic niches were filled (Simpson, 1944, 1953). Recovering the signature of a decrease in the rate of morphological evolution following the invasion of the pelagos, however, could result from a decline in this rate that actually began at the root of the OPM phylogeny and continued through the ecological transition. This might be misinterpreted as a slow down in trophic diversification resulting from the filling of niches in the novel, pelagic adaptive zone (i.e., adaptive radiation). Therefore, we first compared the fit of 1-Rate Brownian motion (BM) models of character evolution for our six traits across the entire OPM clade to a 2-Rate model with a different rate after the benthic to pelagic shift. We used the package *OUwie* v.1.40 (Beaulieu and O'Meara, 2014) and performed these analyses on both the Cyt*b* and Rag1 MCC phylogenies, as well as, across 100 randomly subsampled trees from our distribution of 900 post-burn in trees.

After gauging the support for an increased rate of trait evolution in the clade arising after the pelagic transition, we then used a separate model fitting analysis to assess support for a slow down in this rate through time within this clade. We compared the fit of an 'Early Burst' (EB) (Blomberg et al., 2003; Harmon et al., 2010) model to a BM model of character evolution in this clade for each of our six morphological traits. The EB model is simply a modified version of the BM model that contains an additional parameter, α, which characterizes the exponential rate of increase (positive values) or decrease (negative values) through time of the BM rate parameter (Blomberg et al., 2003; Harmon et al., 2010). Under a BM model $\alpha = 0$. We conducted this model fit comparison across the Cyt*b* and Rag1 MCC phylogenies and our two sets of 100 subsampled post-burn in trees. We ran 1000 iterations of each maximum likelihood search for each model to ensure they did not get stuck at local optima. Model fit between the BM and EB

models was compared using AICc values calculated by *geiger* v.1.99-3.1 (Harmon et al., 2008b).

Finally, we performed a novel analysis in order to examine the effect of the magnitude of the α estimate on the exponential rate of change in the evolution of a trait following the pelagic transition. We performed this analysis on the four traits displaying the strongest evidence for a correlation with JPA (AM1 mass, AM2 mass, lower jaw length, and kinethmoid length) (see Results) using the more thoroughly sampled Cyt*b* MCC phylogeny. These four traits also received the most support for undergoing an early burst of evolution denoted by negative estimates of α on the Cyt*b* MCC tree (see Results). A negative α estimate suggests a decrease in the rate of character evolution towards the present, but the magnitude of this difference in rate relative to the base of the clade is not readily apparent by examining the estimate of α alone. In order to account for the effect of model uncertainty on parameter estimation of α, we first calculated the model average estimate of α for our four traits (Burnham and Anderson, 2002). Using the EB model and the model average parameter estimate of α for a given character, we then transformed the branching times of the Cyt*b* MCC phylogeny using the *transform.phylo* function in *geiger* v.1.99-3.1 (Harmon et al., 2008b). We standardized the untransformed and transformed branching times by dividing by the tree height for each set in order to set the overall tree height equal to 1. Finally, we examined plots of the transformed branching times relative to the untransformed branching times as a function of distance from the root. This allowed us to calculate how much slower the character was evolving through time relative to immediately following the pelagic transition given a negative model average estimate of α.

Results

JPA and foraging height in the water column

The 15 species of OPM cyprinids co-occurring in the Little River displayed significant variation in both JPA and preferred foraging height in the water column (Table 3.1 and Figure 3.2). Species such as *Lythrurus lirus*, *Notropis telescopus*, and *Luxilus coccogenis* were consistently observed foraging close to the water's surface. These taxa had some of the lowest values for JPA. Conversely, we observed species such as *Phenacobius uranops*, *Campostoma oligolepis*, and *Hybopsis amblops* consistently foraging from, or close to, the substrate. These species are at the high end of the range of JPA values. Along with the species inhabiting the extremes of the habitat space, there were species such as *Cyprinella spiloptera*, *Cyprinella galactura*, and *Luxilus chrysocephalus* that preferred the middle sections of the water column on average and tended to display intermediate JPA values. The correlation between PICs of mean JPA and median foraging height score for a species was strongly negative and highly significant based on both MCC phylogenies, as well as, the distribution of post-burn in trees (Table 3.2).

Functional morphological correlates of JPA

AM1 mass, lower jaw length, and kinethmoid length were consistently recovered as being significantly correlated with JPA after phylogenetic correction based on both Cyt*b* and Rag1 (Table 3.2). AM1 displayed a strong positive correlation with JPA across analyses, while lower jaw length and kinethmoid length displayed strong negative correlations across analyses. AM2 mass and AM3 mass received varying degrees of support for a correlation with JPA depending on the locus and analysis (Table 3.2). Overall, a correlation between AP length and JPA received the least amount of support.

Rates of morphological evolution in correlates of JPA

Our 1-Rate versus 2-Rate BM model fitting analyses supported a change in the rate of evolution after the transition to pelagic habitats and foraging modes in a number of the morphological traits examined (Table 3.3). The 2-Rate BM model fit significantly better to the evolution of AM1 mass across both MCC phylogenies, with ΔAICc scores of 10.80 (Cyt*b*) and 7.89 (Rag1) relative to a 1-Rate model. Furthermore, the MCC relative rate estimates suggest that the greatest increase in the rate of evolution after the benthic to pelagic shift occurred in AM1 mass. This rate was estimated to be 3.80 times faster after the shift than before based on the Cyt*b* MCC tree and 3.28 times faster based on the Rag1 MCC phylogeny. Similarly, kinethmoid length received strong support as evolving 3.12 times faster after the habitat shift based on the Cyt*b* MCC tree. The 2-Rate BM model also provided the better fit to the evolution of lower jaw length on both MCC phylogenies. However, the relative rate of lower jaw length evolution was estimated to be slower after the benthic to pelagic transition based on both MCC trees. The results of this analysis varied between the two MCC phylogenies for the remaining three characters (Table 3.3).

Model fitting of 1-Rate versus 2-Rate BM models across our post-burn in distributions of trees displayed several of notable patterns. With the exception of AM3 mass and lower jaw length, the remaining four characters were best modeled as having a separate and faster rate of evolution after the pelagic shift across a large majority of the Cyt*b* post-burn in phylogenies (Table 3.3). Furthermore, although lower jaw length received strong support for a 2-Rate model with a slower rate after the ecological transition using the MCC trees, 76% of the post-burn in Cyt*b* phylogenies were best-fit by a 1-Rate model for this trait (Table 3.3). Conversely, 91% of the post-burn in Rag1 trees also supported a 2-Rate model with a slower rate of lower jaw length evolution before the pelagic shift. Otherwise, the Rag1 post-burn in results were largely variable depending on the trait. However, of the six morphological characters, AM1 mass was most consistently best-fit by a model with a separate and higher rate of evolution after the habitat transition across the Rag1 post-burn in distribution (Table 3.3).

The EB model provided a slightly better fit to the rate of evolution of AM1 mass, AM2 mass, and lower jaw length after the invasion of pelagic habitats based on the Cyt*b* MCC tree, while BM fit slightly better for the remaining three characters (Table 3.4). Notably, however, the maximum likelihood EB model for all six traits included a

negative α estimate based on this phylogeny. Furthermore, the α estimate was lowest for the four characters (AM1 mass, AM2 mass, lower jaw length, and kinethmoid length) that were most significantly correlated with JPA (Tables 3.3 and 3.4). With the exception of lower jaw length, the remaining six characters were almost always better fit by a BM model across the post-burn in distribution of Cyt*b* trees. However, the range of ΔAICc between EB and BM models was generally small, and the estimates of α were consistently negative for the best-fit EB model for the four aforementioned strongest correlates of JPA (Table 3.4). AM1 mass and lower jaw length received a small amount support for an EB model with a negative α as the best-fit model across the Rag1 phylogenies (Table 3.4). Otherwise, BM was generally favored on the less thoroughly sampled Rag1 dataset.

Our final analyses allowed us to graphically illustrate the relationship between a given negative α estimate for a trait and the relative decrease in the rate of evolution of that trait through time in the clade arising after OPM cyprinids shifted to pelagic habitats and foraging modes. The model average estimates of α for the maximum likelihood EB model of trait evolution based on the Cyt*b* MCC tree were -0.06 for AM1 and AM2 mass, -0.03 for lower jaw length, and -0.01 for kinethmoid length. Our plots, which are based on tree transformations using these model average α estimates, display the relative exponential decline in the rate of evolution of these four characters towards the present in the pelagic clade (Figures 3.1 and 3.3). These plots, coupled with results from model fitting analyses on the Cyt*b* MCC tree, suggest that AM1 and AM2 mass are evolving approximately 3 times slower towards the present relative to the period immediately following the OPM transition into pelagic habitats (Figures 3.3A,B). Lower jaw length is estimated to be evolving just under 2 times slower towards the present in this clade (Figure 3.3C). The low model average estimate of α for kinethmoid length is due in part to this being the only one of these four traits whose evolution across the Cyt*b* MCC tree was slightly better modeled by a BM process. This is reflected in the kinethmoid plot as a minimal decrease in its rate of evolution towards the present (Figure 3.3D).

Discussion

The theory of adaptive radiation predicts rapid and simultaneous generation of species diversity and phenotypic variation as organisms evolve traits that allow for exploitation of different niches (Simpson, 1944, 1953; Schluter, 2000). In an earlier study on the evolutionary ecology of OPM minnows (Hollingsworth et al., 2013), we demonstrated a burst of phylogenetic diversification after this clade underwent a major shift from benthic habitats to an open pelagic adaptive zone. This benthic to pelagic axis is often cited to drive microevolutionary divergence in lentic fishes (Robinson and Wilson, 1994; Schluter, 2000; Willacker et al., 2010; Hulsey et al., 2013). The results from this study, however, provide further evidence that this ecological axis can also influence macroevolutionary patterns and even drive adaptive radiations in more diverse clades of lotic fishes.

We conducted the first explicit tests of a relationship between JPA and relative foraging height in the water column in one of the most species rich OPM cyprinid

communities in North America (Etnier and Starnes, 1993). We found a strong negative correlation between this eco-morphological trait and resource partitioning in the form of vertically segregated foraging zones from the benthos to the water's surface. Species with high JPA consistently foraged from or near the benthos, those with intermediate JPA values foraged in the middle sections of the water column, and those with terminal jaws and low JPA specialized on feeding at or near the water's surface. Interestingly, a small number of lineages diverging after the benthic to pelagic shift, such as *Hybopsis amblops* in this community, have returned to feeding predominately from the benthos and can be inferred to have revolved high JPA (Figure 3.1 and Table 3.1) (Hollingsworth et al., 2013). As such, variation in JPA likely facilitates the coexistence of cyprinid fishes in diverse lotic communities by allowing for variable exploitation of the three-dimensional habitat space. Therefore, we should also expect JPA to have experienced a period of explosive diversification coincident with the shift into pelagic habitats if OPM cyprinids have indeed adaptively radiated along a benthic/pelagic habitat axis.

Although the evolution of jaw protrusion and JPA are clearly important eco-morphological traits that drive the diversification of fishes (Schaffer and Rosen, 1961; Lauder, 1982; Ferry-Graham and Lauder, 2001; Konstantinidis and Harris, 2011), the contribution of individual muscoskeletal structures operating during cypriniform jaw protrusion and how they influence JPA are uncertain (Alexander, 1966; Ballintijn et al., 1972; Motta, 1984). In this study we identify a number of traits that are strongly evolutionarily correlated with JPA, including; AM1 mass, AM2 mass, lower jaw length, and kinethmoid length. The strong positive correlation between AM1 mass and JPA corroborates past hypotheses that the adduction of the maxilla by AM1 enhances ventrally directed jaw protrusion, or higher values of JPA, in cypriniform fishes (Ballintijn et al., 1972; Hernandez et al., 2007). AM2, conversely, is rarely active during cypriniform jaw movement (Ballintijn et al., 1972), and the strong relationship between the mass of this muscle and JPA is more difficult to explain. However, Hulsey and Hollingsworth (2011) recovered evidence for correlated evolution between AM1 and AM2 mass, and given their common developmental origin (Hatta et al., 1990; Hernandez et al., 2005), it is possible this is due to auto-correlation between these two muscles that lacks a functional explanation. AM3 is known to function as the main lower jaw adductor (Ballintijn et al., 1972) and showed no significant correlation with JPA. This result was somewhat surprising but could be due to a functional trade-off between force modified and speed modified mechanisms of lower jaw closing (Wainwright et al., 2004; Collar et al., 2008).

Lower jaw length displayed the strongest correlation with JPA over all the characters analyzed in this study. This relationship is qualitatively obvious and is likely due to the general configuration of the mechanism of cyprinid jaw protrusion. Due to the dorsoventral flipping of the kinethmoid pushing the premaxilla in a more subterminal direction (Alexander, 1966; Ballintijn et al., 1972; Hernandez et al., 2007; Staab et al., 2012), benthic species with higher JPA have more overhanging snouts. This leads to a spatiostructural trade-off between JPA and lower jaw length. Species with more terminal mouths and lower JPA can have longer lower jaws. The kinethmoid was also strongly correlated with JPA. This relationship suggests a longer kinethmoid might be able to

rotate relatively further dorsoventrally during jaw protrusion, ultimately pushing the premaxilla in a more subterminal direction. The length of the AP was only marginally significantly correlated with JPA. The length of the AP of the premaxilla has been shown to be correlated with the distance of jaw protrusion in cichlid fishes (Hulsey et al., 2010), and although we did not quantify protrusion distance, species with higher JPA have qualitatively greater protrusion distance. This could explain this marginal correlation, but would need to be verified by quantifying protrusion distance across these cyprinid species.

Our model fitting analyses revealed a number of patterns of variation in the rates of morphological evolution before and after the pelagic transition that are consistent with an adaptive radiation in OPM cyprinids mediated by diversification of JPA. The two traits that displayed the strongest correlations to JPA, AM1 mass and lower jaw length, displayed the most consistent results across model fitting analyses. From a functional standpoint, the high correlation between AM1 mass and JPA recovered in this study and the role of this muscle in driving benthicly directed jaw protrusion in cyprinids (Ballintjin et al., 1972), together suggest that the evolution this muscle may have the most direct impact on divergence in JPA. AM1 was best modeled as evolving faster (higher relative BM rate) after the transition into pelagic habitats and also experiencing a slowing of this rate (best fit by EB model) towards the present. This is consistent with the expectations of phenotypic evolution of relevant traits during an adaptive radiation (Simpson, 1944, 1953). Lower jaw length on the other hand received very strong support for having a slower rate of evolution before the benthic to pelagic transition, but also received the most support for a slow down in this rate after the habitat shift. This pattern could be attributable to two causes. The estimated slower rate before the pelagic shift could be accurate and the EB model is providing the better fit to lower jaw length evolution after the shift because the rate has continued to slow throughout the OPM phylogeny. If, however, lower jaw length evolution experienced a dramatic slow down in rate after the pelagic shift and towards the present, then our fitting of 1-Rate versus 2-Rate BM models to the evolution of this trait could be misleading. For instance, the evolution of lower jaw length after the benthic to pelagic transition was best fit by an EB model with a relatively large negative estimate of α based on the Cyt*b* MCC and post-burn in phylogenies (Table 3.4). A rapid decline in the rate of evolution in this trait through time could result in very slow rates at the tips. When fitting a BM model to the evolution of lower jaw length over this portion of the OPM phylogeny, this dramatic slow down could result in a lower average BM rate estimate and give the false appearance of a slower rate of evolution in this clade. This issue could be addressed by the development of methods that simultaneously fit BM and EB models to different segments of a phylogeny under consideration.

The remaining four morphological correlates of JPA were much less consistently best fit by the same models across loci and analyses. Furthermore, in the BM versus EB model fitting analyses we seemed to lack the power necessary to reject BM in many cases (Slater and Pennell, *in press*). We saw evidence for this in our results in two ways. First, the less thoroughly sampled Rag1 dataset, which contained 16 less species from the clade arising after the pelagic shift, was nearly always better fit by a BM model after the habitat

transition relative to the Cyt*b* dataset. Therefore, we also conducted a pilot analysis with a Cyt*b* dataset missing the same 16 OPM taxa. Using this less-inclusive Cyt*b* dataset resulted in a notable loss of support for the EB model of trait evolution after the habitat shift relative to the more complete Cyt*b* dataset used in this study. There are approximately 190 species that comprise the clade arising after the pelagic shift (Hollingsworth et al., 2013). As such, these results illustrate that a mere 8% decrease in taxon sampling has a large affect on the power to reject a BM model in favor of an EB model in this clade. Another factor that can bias these analyses away from detecting an early burst of morphological evolution is convergence towards the tips (Slater and Pennell, *in press*). Quantifying convergence and accounting for it in such analyses, however, warrants further research.

In this study, we demonstrate a solid link between an eco-morphological trait, JPA, and resource partitioning in the form of vertically stratified foraging zones within a diverse assemblage of OPM cyprinid fishes. We then present evidence for a historical burst of diversification in JPA after this clade shifted from benthic habitats and foraging modes into a previously unoccupied, and further subdivisible, pelagic zone. Along with results from Hollingsworth et al. (2013), which suggest that a period of rapid cladogenesis followed this ecological transition, our findings provide strong support for an adaptive radiation of cyprinids along a three-dimensional benthic to pelagic axis. Taken together, we feel as though this set of studies represents one of the most robust tests for an adaptive radiation yet to be conducted and highlights the role of this ecological axis as a ubiquitous generator of biodiversity.

References

Alexander, R.M., 1966. The functions and mechanisms of the protrusible upper jaw of two species of cyprinid fish. *J. Zool. Lond.* 149, 288-296.

Alexander, R.M., 1967. The functions and mechanisms of the protrusible jaws of some acanthopterygian fishes. *J. Zool. Lond.* 151, 233-255.

Baker, J.A., Ross, S.T., 1981. Spatial and temporal resource utilization by southeastern cyprinids. *Copeia* 1981, 178-189.

Ballintijn, C.M., Van Den Burg, A., Egberink, B.P., 1972. An electromyographic study of the adductor mandibulae complex for a free-swimming carp *Cyprinus carpio* during feeding. *J. Exp. Biol.* 57, 261-283.

Beaulieu, J.M., O'Meara, B.C., 2014. OUwie: Analysis of evolutionary rates in an OU framework. R package version 1.40.

Blomberg, S.P., Garland, T., Ives, A.R., 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* 57, 717-745.

Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference : a practical information-theoretic approach. Springer, New York, NY.

Cochran-Biederman, J.L., Winemiller, K.O., 2010. Relationships among habitat, ecomorphology and diets of cichlids in the Bladen River, Belize. *Environ. Biol. Fish.* 88, 143–152.

Collar, D.C., Wainwright, P.C., Alfaro, M.E., 2008. Integrated diversification of locomotion and feeding in labrid fishes. *Biol. Lett*. 4, 84-86.

DeVries, P.J., 1988. Stratification of fruit-feeding nymphalid butterflies in a Costa Rican rainforest. *J. Res. Lepidoptera* 26, 98-108.

DeVries, P.J., Alexander, L.G., Chacon, I.A., Fordyce, J.A., 2012 Similarity and difference among rainforest fruit-feeding butterfly communities in Central and South America. *J. Anim. Ecol.* 81, 472-482.

Dingerkus, G., Uhler, L.H., 1977. Enzyme clearing of Alcian blue stained whole vetebrates for demonstration of cartilage. *Stain Technol*. 52, 229-232.

Dumont, E.R., Dávalos, L.M., Goldberg, A., Santana, S.E., Rex, K., Voigt, C.C., 2012. Morphological innovation, diversification, and invasion of a new adaptive zone**.** *Proc. R. Soc. B* 279, 1797-1805.

Etnier, D.A., Starnes, W.C., 1993. The fishes of Tennessee. University of Tennessee Press, Knoxville, TN.

Felsenstein, J., 1985. Phylogenies and the comparative method. *Am. Nat.* 125, 1-15.

Ferry-Graham, L.A., Lauder, G.V., 2001. Aquatic prey capture in ray-finned fishes: a century of progress and new directions. *J. Morphol.* 248, 99-119**.**

Fordyce, J.A., 2010. Host shifts and evolutionary radiations of butterflies. *Proc. R. Soc. B* 277, 3735-3743.

Garland Jr., T., Harvey, P.H., Ives, A.R., 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Syst. Biol.* 41, 18-32.

Gatz Jr., A.J., 1979. Ecological morphology of freshwater stream fishes. *Tulane Stud. Zool. Bot.* 21, 91-124.

Gorman, O.T., 1988a. The dynamics of habitat use in a guild of Ozark minnows. *Ecol. Monogr.* 58, 1-18.

Gorman, O.T., 1988b. An experimental study of habitat use in an assemblage of Ozark minnows. *Ecology* 69, 1239-1250.

Harmon, L.J., Melville, J., Larson, A., Losos, J.B., 2008a. The role of geography and ecological opportunity in the diversification of day geckos (*Phelsuma*). *Syst. Biol.* 57, 562-573.

Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E., Challenger, W., 2008b. GEIGER: investigation evolutionary radiations. *Bioinformatics* 24, 129-131.

Harmon, L.J., Losos, J.B., Davies, T.J., Gillespie, R.G., Gittleman, J.L., Jennings, W.B., Kozark, K.H., McPeek, M.A., Moreno-Roark, F., Near, T.J., Purvis, A., Ricklefs, R.E., Schluter, D., Schulte II, J.A., Seehausen, O., Sidlauskas, B.L., Torres-Carvajal, O., Weir, J.T., Mooers, A.Ø., 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64, 2385-2396.

Hatta, K., Schilling, T.F., Bremiller, R.A., Kimmel, C.B., 1990. Specification of jaw muscle identity in zebrafish – correlation with engrailed homeoprotein expression. *Science* 250, 802–805.

Hernandez, L.P., Patterson, S.E., Devoto, S.H., 2005. The development of muscle fiber type identity in zebrafish cranial muscles. *Anat. Embryol.* 209, 323–334.

Hernandez, L.P., Bird, N.C., Staab, K.L., 2007. Using zebrafish to investigate cypriniform evolutionary novelties: functional development and evolutionary diversification of the kinethmoid. *J. Exp. Zool.* (*Mol. Dev. Evol.*) 308B, 625-641.

Hollingsworth Jr., P.R., Simons, A.M., Fordyce, J.A., Hulsey, C.D., 2013. Explosive diversification following a benthic to pelagic shift in freshwater fishes. *BMC Evol. Biol.* 13, 272.

Hulsey, C.D., Hollingsworth Jr., P.R., 2011. Do constructional constraints influence cyprinid (Cyprinidae: Leuciscinae) craniofacial coevolution? *Biol. J. Linn. Soc.* 103, 136-146.

Hulsey, C.D., Hollingsworth Jr., P.R., Holzman, R., 2010. Co-evolution of the premaxilla and jaw protrusion in cichlid fishes. *Biol. J. Linn. Soc.* 100, 619-629.

Hulsey, C.D., Roberts, R.J., Loh, Y.H.E., Rupp, M.F., Streelman, J.T., 2013. Lake Malawi cichlid evolution along a benthic/limnetic axis. *Ecol. Evol.* 3, 2262-2272.

Konstantinidis, P., Harris, M.P., 2011. Same but different: ontogeny and evolution of the *Musculus adductor mandibulae* in the Tetraodontiformes. *J. Exp. Zool. Part B* 1, 10-20.

Lauder, G.V., 1982. Patterns of evolution in the feeding mechanism of actinopterygian fishes. *Amer. Zool.* 22, 275–285.

Losos, J.B., 2010. Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am. Nat.* 175, 623-639.

Mendelson, J., 1975. Feeding relationships among species of *Notropis* (Pisces: Cyprinidae) in a Wisconsin stream. *Ecol. Monogr.* 45, 199-230.

Motta, P.J., 1984. Mechanics and functions of jaw protrusion in teleost fishes: a review. *Copeia* 1984, 1-18.

O'Meara, B.C., Ané, C., Sanderson, M.J., Wainwright, P.C., 2006. Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60, 922-933.

Osse, J.W.M., 1969. Functional morphology of the head of the perch (Perca fluviatilis): an electromyographical study. *Neth. J. Zool.* 19, 289–392.

Page, L.M., Burr, B.M., 2011. Peterson field guide to freshwater fishes, second edition. Houghton Mifflin, Boston, MA.

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

R Development Core Team, 2013. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. [http://www.R-project.org/]

Revell, L.J., 2009. Size-correction and principal components for interspecific comparative studies. *Evolution* 63, 3258–3268.

Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217-223.

Robinson, B.W., Wilson, D.S., 1994. Character release and displacement in fishes: a neglected literature. *Am. Nat.* 144, 596-627.

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

Schaeffer, B., Rosen, D.E., 1961. Major adaptive levels in the evolution of the actinopterygian feeding mechanism. *Amer. Zool.* 1, 187–204.

Simpson, G.G., 1944. Tempo and mode in evolution. Columbia University Press, New York, NY.

Simpson, G.G., 1953. The major features of evolution. Columbia University Press, New York, NY.

Slater, G.J., Pennell, M.W., *in press*. Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. *Syst. Biol.*

Staab, K.L., Ferry, L.A., Hernandez, L.P., 2012. Comparative kinematics of cypriniform premaxillary protrusion. *Zoology* 115, 65-77.

Yoder, J.B., Clancey, E., Des Roches, S., Eastman, J.M., Gentry, L., Godsoe, W., Hagey, T.J., Jochimsen, D., Oswald, B.P., Robertson, J., Sarver, B.A.J., Schenk, J.J., Spear, S.F., Harmon, L.J., 2010. Ecological opportunity and the origin of adaptive radiations. *J. Evol. Biol.* 23, 1581-1596.

Wainwright, P.C., Bellwood, D.R., Westneat, M.W., Grubich, J.R., Hoey, A.S., 2004 A functional morphospace for the skull of labrid fishes: patterns of diversity in a complex biomechanical system. *Biol. J. Linn. Soc. Lond.* 82, 1–25.

Westneat, M.W., 2004. Evolution of levers and linkages in the feeding mechanisms of fishes. *Integr. Comp. Biol.* 44, 378–389.

Willacker, J.J., von Hippel, F.A., Wilton, P.R., Walton, K.M., 2010. Classification of threespine stickleback along the benthic-limnetic axis. *Biol. J. Linn. Soc. Lond.* 101, 595-608.

CONCLUSION

This dissertation explored the evolutionary ecology of North America's most diverse clade of endemic freshwater fishes, OPM cyprinids, as they have diversified along a benthic/pelagic habitat axis. This ecological axis is often cited to drive microevolutionary divergence in lentic fishes, but its influence on the macroevolution of diverse clades of lotic fishes has until now remained unclear. In Chapter I, I used state of the art phylogenetic methods to resolve long-standing discordance between phylogenetic hypotheses for this group of fishes. The strongly supported general topology recovered in Chapter I suggests that benthic forms dominated the early history of this clade and then a major habitat shift up into the pelagic zone took place. In Chapter II, I recovered evidence for a burst of lineage diversification following this ecological shift and hypothesized that the benthic to pelagic transition represented the invasion of a novel adaptive zone. Based on the theory of adaptive radiation, which predicts rapid and simultaneous lineage and phenotypic divergence in a clade upon the entrance of an open adaptive zone, I designed studies in Chapter III to combine with the results from Chapter II and robustly test for signature of an adaptive radiation in OPM cyprinids. I demonstrate a strong evolutionary relationship between an important eco-morphological trait, JPA, and vertical microhabitat partitioning in a diverse OPM assemblage. I then find substantial evidence for a burst of morphological evolution in a number of individual muscoskeletal trophic characters that are significantly correlated with JPA after the benthic to pelagic shift. Taken together, the results from this dissertation strongly suggest that the shift into pelagic habitats allowed this clade to explosively radiate as lineages evolved variation in trophic morphology to exploit vertically segregated habitats within a previously unoccupied pelagic zone. Therefore, the influence of this habitat axis is clearly not only responsible for driving microevolutionary divergence in lentic fishes, but instead appears to be a ubiquitous generator of biodiversity across disparate freshwater fishes and ecosystems. Furthermore, this dissertation likely provides the most thorough tests of, and evidence for, an adaptive radiation to date.

APPENDIX

Table 1.1. Genetic partition information. The length in base pairs (bp) of each partition analyzed in Chapter I is given. Zic1 does not include first position codon sites. Models chosen by MrModelTest2 for each partition (Nylander, 2004) are abbreviated: GTR (general time reversible), HKY (Hasegawa, Kishino, and Yano), K80 (Kimura 80), SYM (symmetrical model), I (proportion of invariant sites), G (gamma distributed substitution rates). Maximum uncorrected sequence divergence is given for each partition, as well as the North American species displaying this divergence. Abbreviated taxa names are detailed in the legend to Figure 1.1.

Partition	Length (bp)	Maximum sequence divergence	Model
Cytb 1st	380	11.9% *Satr* - *Pcat*	SYM + I + G
Cytb 2nd	380	1.6% *Neff* - *Pmir*	HKY
Cytb 3rd	380	52.5% *Satr* - *Pmir*	GTR + I + G
Enc1	810	3.8% *Satr* - *Edis*	HKY + G
Ptr	699	2.6% *Satr* - *Pcra*	GTR + G
Ryr3	822	3.2% *Satr* - *Pnot*/*Pcat*/*Pmir*/*Pura*	HKY + I
Sh3px3	705	4.3% *Pcat* - *Nleu*	K80 + G
Tbr1	645	2.0% *Satr* - *Elau*	HKY + I
Zic1	572	3.3% *Coli* - *Edis*	HKY + I

Table 2.1. Tests for an early burst of diversification in the pelagic clade. These results represent the analyses carried out on the Cyt*b* and Rag1 MCC gene trees. The "best model" indicates the best-fit model of cladogenesis chosen for this clade based on AIC scores. (MCCR = Monte Carlo constant rates test, TD = tree deviation test, DDL = density dependent logistic, Y2R = Yule 2-rate)

Locus	γ	MCCR *p*-value	TD *p*-value	ΔAIC *p*-value	best model
Cyt*b*	-7.59	< 0.001	< 0.001	< 0.001	DDL
Rag1	-2.55	< 0.058	0.005	0.024	Y2R

Table 3.1. Preferred foraging height and JPA across the Little River OPM assemblage. The median of foraging height scores and mean JPA ± SE for 15 OPM cyprinid species that co-occur in Little River, Blount Co., TN are given. (n) = sample size in numbers of individuals.

Species	Median Height (n)	Mean JPA ± SE (n)
Lythrurus lirus	4 (24)	7.50 ± 0.38 (8)
Notropis telescopus	3.5 (50)	10.30 ± 0.32 (23)
Luxilus coccogenis	3 (50)	9.83 ± 0.35 (23)
Notropis micropteryx	3 (50)	13.64 ± 0.39 (25)
Notropis photogenis	3 (25)	8.30 ± 0.28 (20)
Cyprinella spiloptera	3 (13)	12.14 ± 0.34 (7)
Luxilus chrysocephalus	2 (50)	9.64 ± 0.31 (25)
Cyprinella galactura	2 (50)	15.35 ± 0.70 (23)
Notropis leuciodus	2 (50)	13.19 ± 0.40 (21)
Notropis volucellus	2 (50)	19.46 ± 0.46 (13)
Erimystax insignis	2 (5)	24.63 ± 0.67 (19)
Nocomis micropogon	1 (50)	20.00 ± 0.47 (10)
Hybopsis amblops	1 (50)	31.20 ± 0.39 (25)
Campostoma oligolepis	1 (50)	27.33 ± 0.29 (30)
Phenacobius uranops	1 (36)	36.40 ± 0.40 (10)

Table 3.2. Correlation coefficients, r, of PICs between JPA and habitat use and JPA and individual morphological characters. MCC r is from the Cyt*b* and Rag1 MCC phylogenies. Median$_{900}$ r is the median correlation coefficient calculated across our 900 subsampled trees and (95% CR) gives the 95% central range of this distribution. The associated *p* values for MCC r and median$_{900}$ r are given below each. Habitat represents the preferred foraging height in the water column (see Table 1). AM1, AM2, and AM3 are the three AM muscle masses. LJ is the length of the lower jaw. AP is the length of ascending process of the premaxilla. Kin represents the length of the kinethmoid.

	MCC r	**median$_{900}$ r (95% CR)**		**MCC r**	**median$_{900}$ r (95% CR)**
Cyt*b*			**Rag1**		
Habitat	-0.72	-0.73 (-0.76, -0.69)	Habitat	-0.80	-0.80 (-0.85, -0.75)
	$p < 0.005$	$p < 0.005$		$p < 0.005$	$p < 0.005$
AM1	0.61	0.63 (0.60, 0.66)	AM1	0.38	0.47 (0.15, 0.58)
	$p < 0.005$	$p < 0.005$		$p = 0.014$	$p < 0.005$
AM2	0.49	0.50 (0.47, 0.53)	AM2	0.26	0.34 (0.10, 0.44)
	$p < 0.005$	$p < 0.005$		$p = 0.101$	$p = 0.030$
AM3	-0.24	-0.27 (-0.33, -0.22)	AM3	-0.37	-0.37 (-0.44, -0.30)
	$p = 0.131$	$p = 0.088$		$p = 0.017$	$p = 0.017$
LJ	-0.68	-0.69 (-0.71, -0.65)	LJ	-0.72	-0.69 (-0.73, -0.64)
	$p < 0.005$	$p < 0.005$		$p < 0.005$	$p < 0.005$
AP	0.30	0.33 (0.28, 0.37)	AP	0.10	0.20 (-0.03, 0.33)
	$p = 0.057$	$p = 0.035$		$p = 0.534$	$p = 0.210$
Kin	0.65	0.63 (0.60, 0.67)	Kin	0.56	0.56 (0.48, 0.63)
	$p < 0.005$	$p < 0.005$		$p < 0.005$	$p < 0.005$

Table 3.3. Comparison of 2-Rate versus 1-Rate Brownian motion models of trait evolution before and after the pelagic transition. MCC best model signifies the best model fit to the two MCC phylogenies and the ΔAICc for the alternative model is given in parentheses. MCC relative rate is the ratio of the estimated rate parameter for the clade arising after the pelagic shift divided by the estimated rate before the habitat transition. The range of difference in AICc values for the 2-Rate AICc minus the 1-Rate AICc is given. The "% 2-Rate favored" column denotes the percentage of 100 subsampled post-burn in phylogenies for which the 2-Rate model had the lower AICc score. The "% higher rate after shift" column contains the percentages of post-burn in trees where the best-fit 2-Rate BM model included a faster rate estimate for that character after the benthic to pelagic transition. AM1, AM2, and AM3 represent the mass of the three AM muscles. LJ denotes lower jaw length. AP is the length of the ascending process of the premaxilla. Kin represents the length of the kinethmoid.

	MCC best model (ΔAICc)	MCC relative rate	Range ΔAICc 2-Rate - 1-Rate	% 2-Rate favored	Range relative rate	% higher rate after shift
Cytb						
AM1	2-Rate (10.80)	3.80	-51.03:2.01	97%	0.01:28.22	67%
AM2	2-Rate (8.90)	3.52	-45.21:1.63	92%	0.02:21.69	67%
AM3	1-Rate (2.39)	0.88	-9.56:7.81	15%	0.41:3.58	70%
LJ	2-Rate (2.82)	0.47	-3.40:2.49	34%	0.10:1.59	17%
AP	1-Rate (0.38)	1.53	-29.03:0.91	96%	4.52E-3:9.98	67%
Kin	2-Rate (7.53)	3.12	-49.40:1.66	94%	2.76E-3:26.13	67%
Rag1						
AM1	2-Rate (7.89)	3.28	-13.40:2.09	44%	0.06:4.77	86%
AM2	1-Rate (0.40)	1.54	-19.60:2.10	18%	0.03:2.77	71%
AM3	1-Rate (1.54)	1.30	-12.23:2.17	25%	0.11:1.99	28%
LJ	2-Rate (14.05)	0.31	-46.89:2.10	91%	0.14:1.00	0%
AP	2-Rate (3.03)	2.22	-13.15:2.10	30%	0.07:4.42	69%
Kin	1-Rate (2.05)	0.91	-2.32:2.16	7%	0.53:1.94	53%

Table 3.4. Comparison of 'Early Burst' versus Brownian motion models of trait evolution after the pelagic transition. The MCC best model is the model with the lowest AICc score for the Cyt*b* and Rag1 MCC trees. The ΔAICc for the alternative model is given in parentheses. MCC α denotes the estimate of the rate-change parameter for the best-fit 'Early Burst' (EB) model for the MCC phylogenies. The range of difference in AICc for the best-fit EB model minus the best-fit Brownian motion (BM) model across our 100 subsampled post-burn in trees is given. The range of α estimates for the best-fit EB model to each of these 100 trees is also given. The "%EB as best model" column is the percentage of post-burn in trees for which the EB model received the lower AICc value. The "% negative α estimate" column contains the percentage of post-burn in phylogenies where the best-fit EB model included a negative α estimate. AM1, AM2, and AM3 represent the mass of the three AM muscles. LJ denotes lower jaw length. AP is the length of the ascending process of the premaxilla. Kin represents the length of the kinethmoid.

	MCC Best Model (ΔAICc)	MCC α	Range EB-BM AICc	% EB as best model	Range α	% negative α
Cytb						
AM1	EB (1.74)	-0.09	-2.23:2.14	16%	-0.15:0.03	89%
AM2	EB (1.53)	-0.08	-1.65:2.14	10%	-0.16:0.09	73%
AM3	BM (1.75)	-0.03	-0.34:2.14	1%	-0.10:0.10	20%
LJ	EB (0.73)	-0.06	-9.94:1.92	91%	-0.28:-0.02	100%
AP	BM (2.11)	-0.01	0.56:2.14	0%	-0.07:0.17	6%
Kin	BM (1.45)	-0.04	-0.13:2.14	1%	-0.10:0.03	85%
Rag1						
AM1	EB (0.38)	-0.10	-3.66:2.17	6%	-0.12:0.22	79%
AM2	BM (2.17)	-1.00E-6	-5.35:2.17	30%	-0.03:0.27	2%
AM3	BM (2.17)	-1.00E-6	-5.99:1.75	50%	0.03:0.29	0%
LJ	BM (1.97)	-0.02	-0.44:2.17	2%	-0.08:0.04	65%
AP	BM (2.17)	-1.00E-6	-3.95:2.17	19%	-3.00E-3:0.25	2%
Kin	BM (2.17)	-1.00E-6	-2.97:2.17	13%	-0.02:0.12	13%

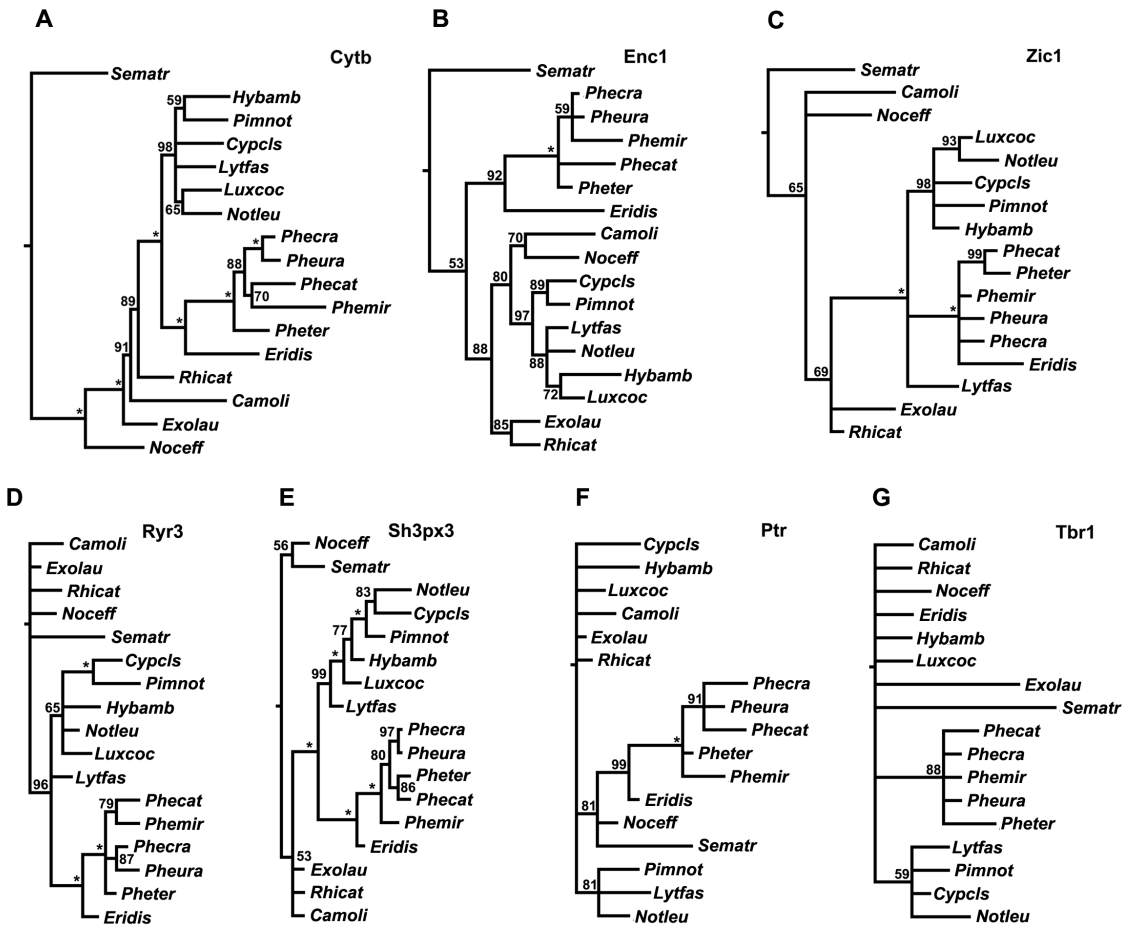

Figure 1.1. Individual gene trees. The 50% majority rule consensus gene trees obtained using MrBayes are presented for the individual loci: A) Cytb B) Enc1 C) Zic1 D) Ryr3 E) Sh3px3 F) Ptr G) Tbr1. Values at nodes are Bayesian posterior support values given in percentages. Asterisks represent 100% posterior support. Taxon names have been abbreviated to save space. *Campostoma oligolepis* (*Camoli*), *Cyprinella callistia* (*Cypcls*), *Erimystax dissimilis* (*Eridis*), *Exoglossum laurae* (*Exolau*), *Hybopsis amblops* (*Hybamb*), *Luxilus coccogenis* (*Luxcoc*), *Lythrurus fasciolaris* (*Lytfas*), *Nocomis effusus* (*Noceff*), *Notropis leuciodus* (*Notleu*), *Phenacobius catostomus* (*Phecat*), *P. crassilabrum* (*Phecra*), *P. mirabilis* (*Phemir*), *P. teretulus* (*Pheter*), *P. uranops* (*Pheura*), *Pimephales notatus* (*Pimnot*), *Rhinichthys cataractae* (*Rhicat*), *Semotilus atromaculatus* (*Sematr*). *Danio rerio* was used to root the phylogenies and this branch was subsequently removed from each gene tree for presentation in this figure.

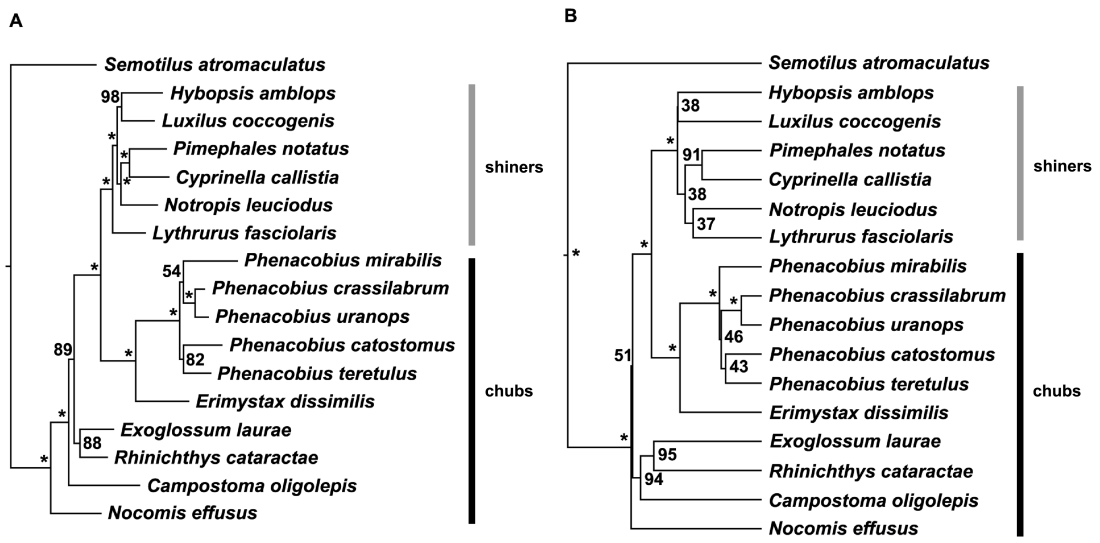

Figure 1.2. Multi-locus phylogenies. A) The 50% majority rule consensus phylogeny is presented from the concatenated Mr. Bayes analysis. B) The maximum clade credibility phylogeny is presented from the *BEAST species tree analysis. Values at nodes are posteriors given in percentages. Asterisks indicate 100% of post burn-in trees contained the clade. “Chub” and “shiner” designations are based on the morphological phylogenetic hypothesis of Mayden (1989). *Danio rerio* was used to root the phylogenies and this branch was subsequently removed from each species tree for presentation in this figure.

Figure 2.1. Concatenated Cyt*b* + Rag1 phylogeny. A), B), and C) correspond to the panels on the smaller version of the complete tree presented to the right. Maximum clade credibility (MCC) phylogeny of OPM cyprinids based on the Cyt*b* + Rag1 analysis. Numbers at nodes represent posterior probability values (pp). Asterisks denote 100% pp.

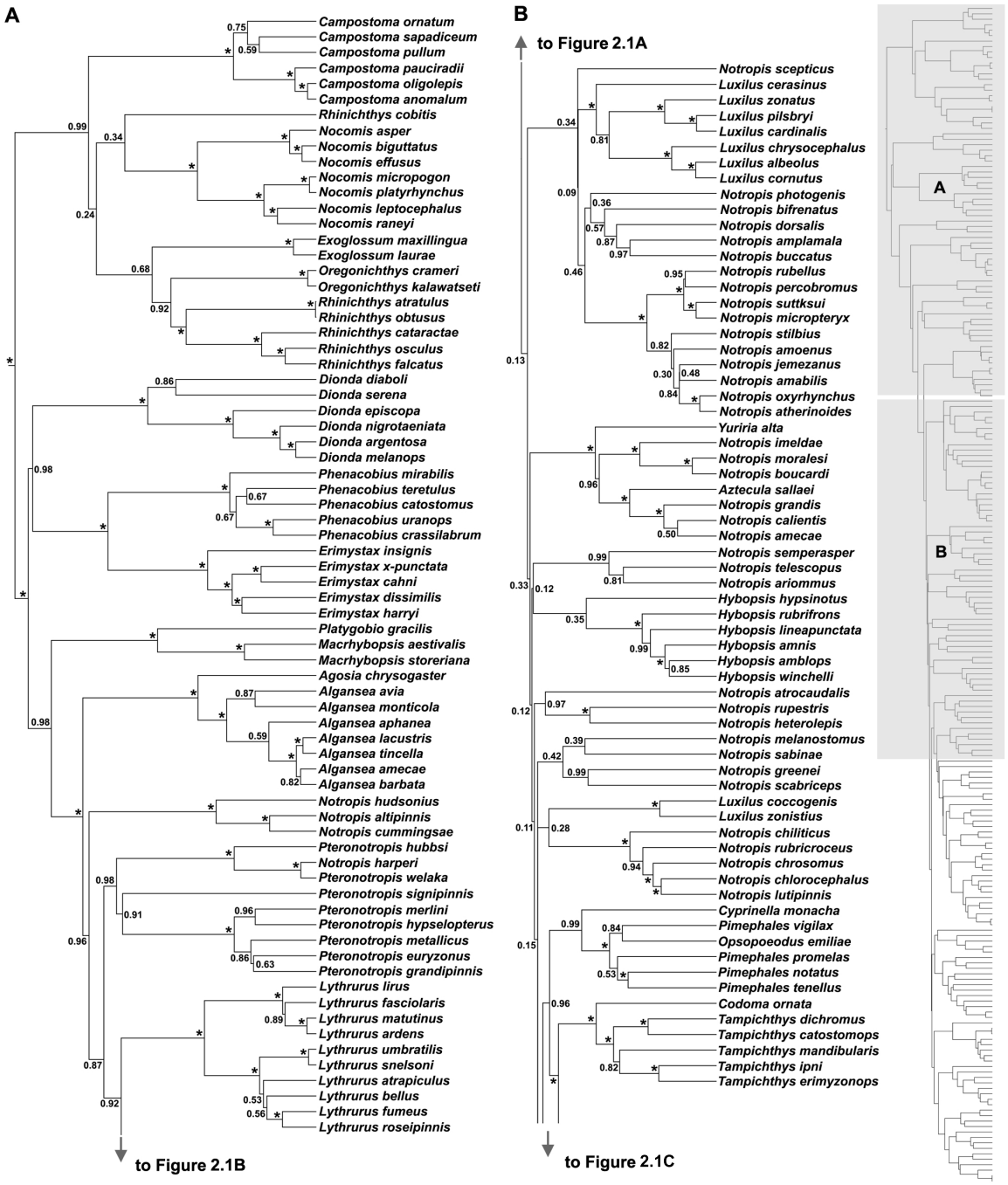

Figure 2.1. continued

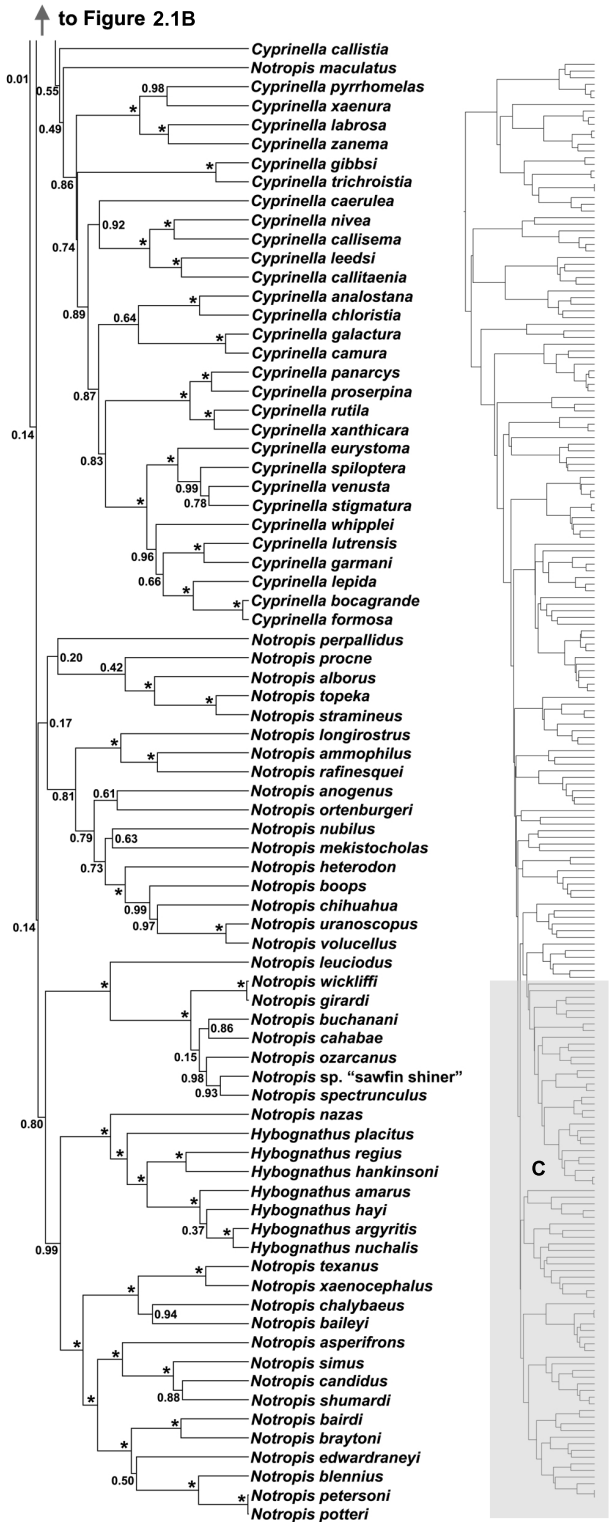

Figure 2.1. continued

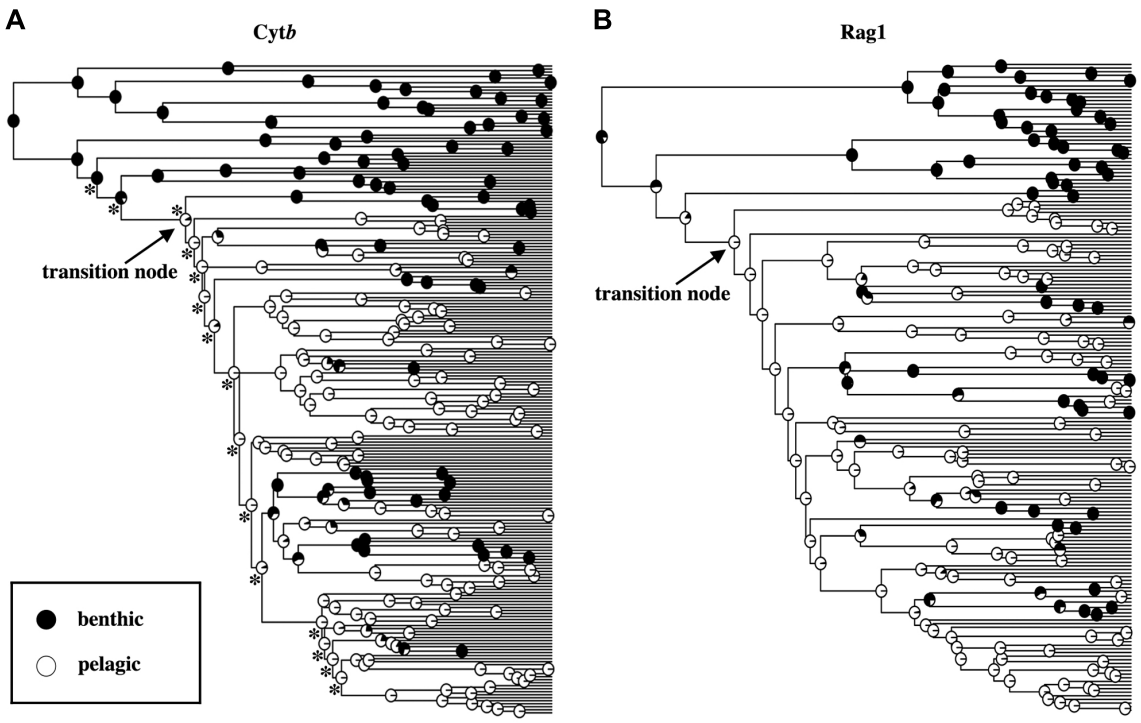

Figure 2.2. Ancestral state reconstructions. A) Benthic/pelagic ancestral state reconstruction on the Cyt*b* MCC gene tree. B) Benthic/pelagic ancestral state reconstruction on the Rag1 MCC gene tree. The 'transition node' indicated by the black arrows represents the first phylogenetically well-supported benthic to pelagic habitat shift. The asterisks in A) denote nodes that were recovered as significantly diverse based on the RC test.

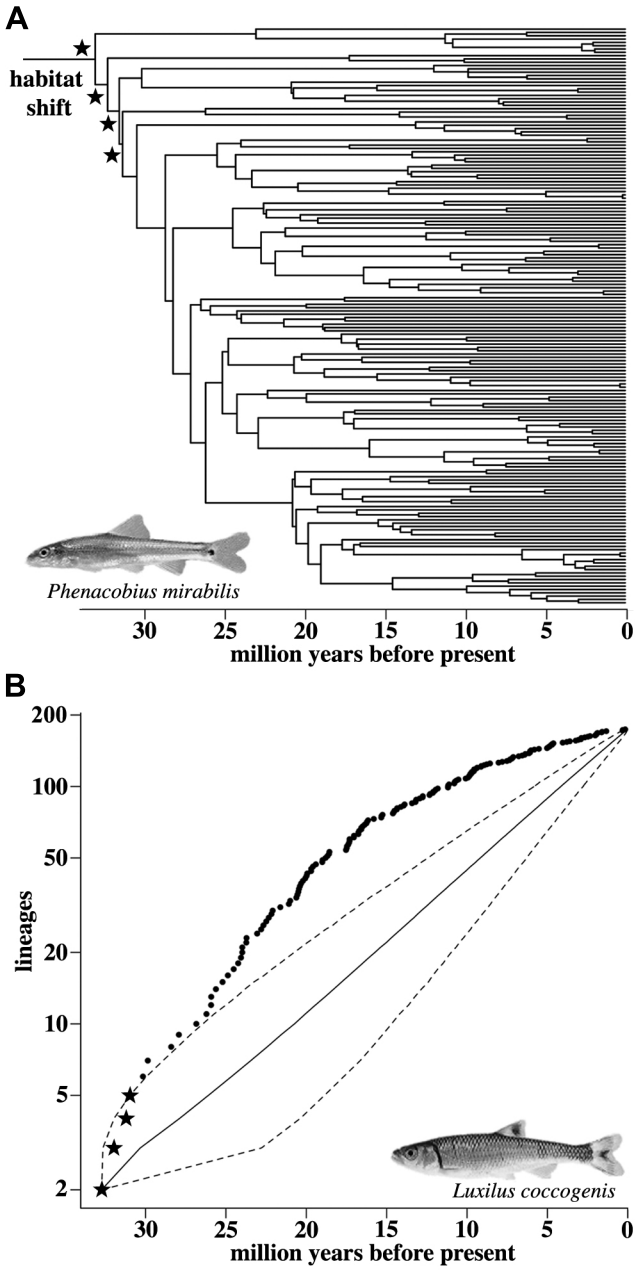

Figure 2.3. Diversification rate analyses for the focal clade. A) The Cyt*b* MCC chronogram. The node labeled 'habitat shift' corresponds to the 'transition node' in (Figure 2.2A). Stars identify subclades that were significant in the PRC analysis at α = 0.1. *Phenacobius mirabilis,* a benthic OPM species, is pictured. B) LTT plot for the Cyt*b* MCC gene tree after the transition node. Black dots indicate the actual number of lineages in our reconstructed tree. The stars correspond to the starred nodes in (Figure 2.3A). The solid black line denotes the mean, and dotted lines the 95% confidence intervals, from 10,000 simulated pure-birth phylogenies. The y-axis is on the log scale. *Luxilus coccogenis*, a pelagic OPM species, is pictured.

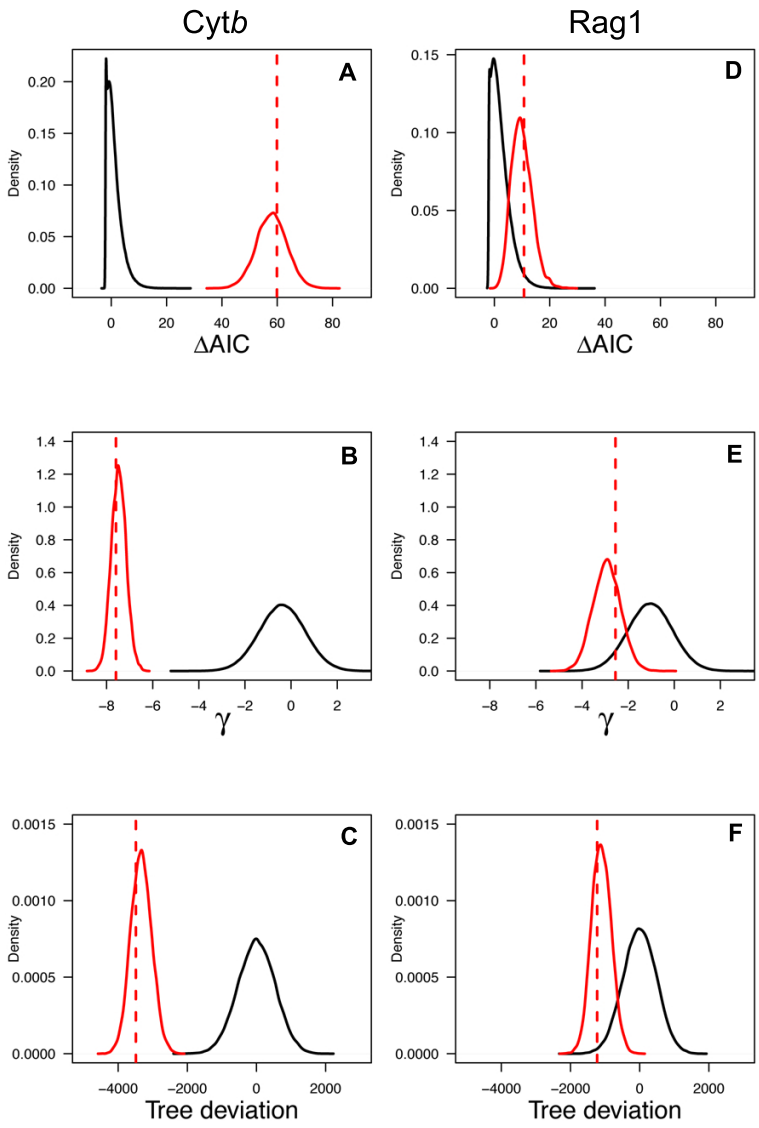

Figure 2.4. Simulated lineage accumulation and $_{\Delta}$AIC values. Density plots for lineage accumulation statistics and ΔAIC values from 1 million simulated pure-birth phylogenies with taxon sampling (black) and 9005 post burn-in trees (red) for Cyt*b* A), B), and C), and Rag1 D), E), and F). Hatched red line indicates values for MCC tree.

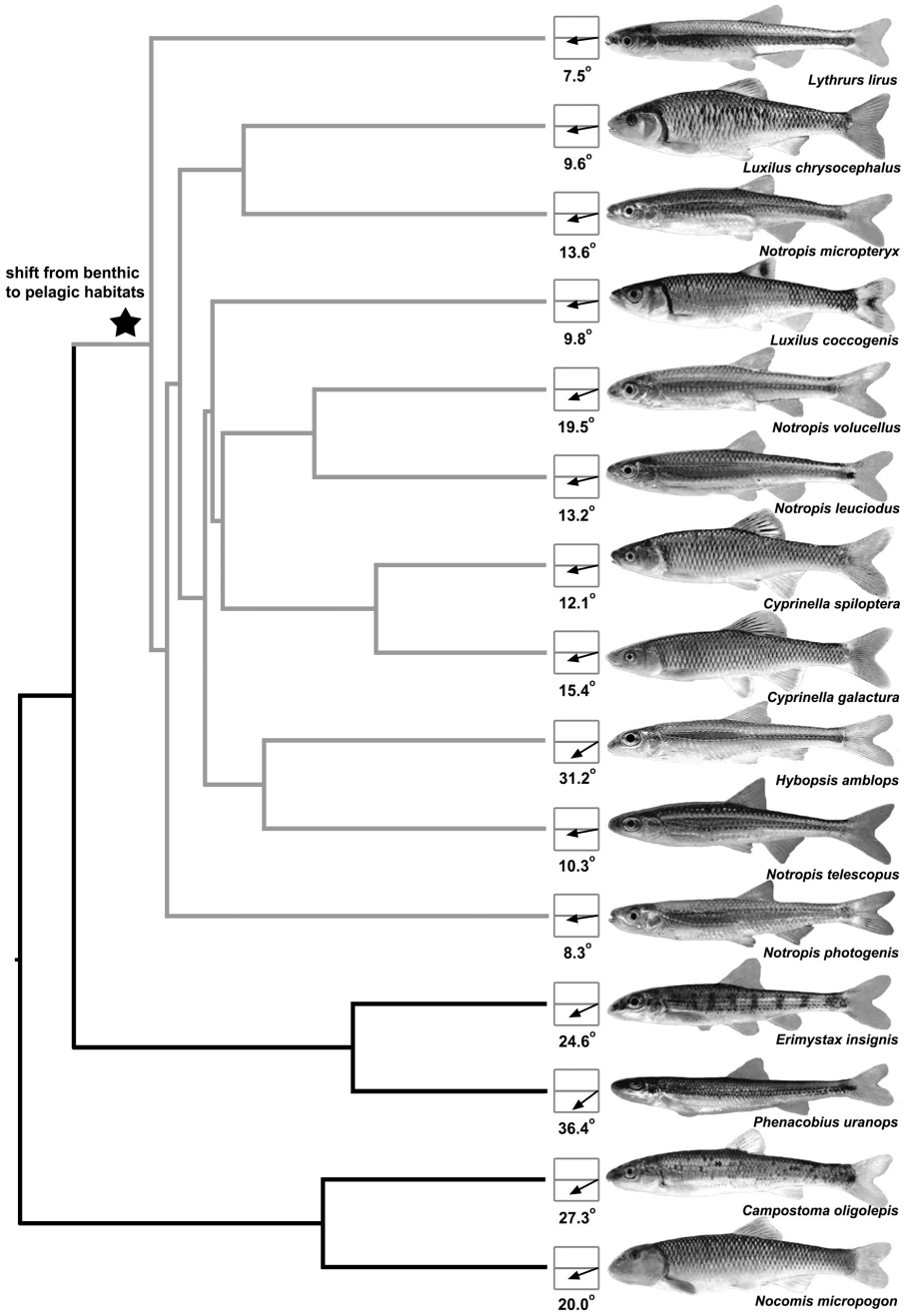

Figure 3.1. Mean JPA and phylogenetic relationships based on Rag1 across the Little River OPM community. The black branches are lineages diverging before the benthic to pelagic shift, and grey branches denote those diverging after. A small number of lineages in the predominately pelagic clade, such as *Hybopsis amblops* from this assemblage, have high JPA values as they have returned to foraging benthicly (Table 3.1). The mean JPA is given for each species above a box illustrating this angle.

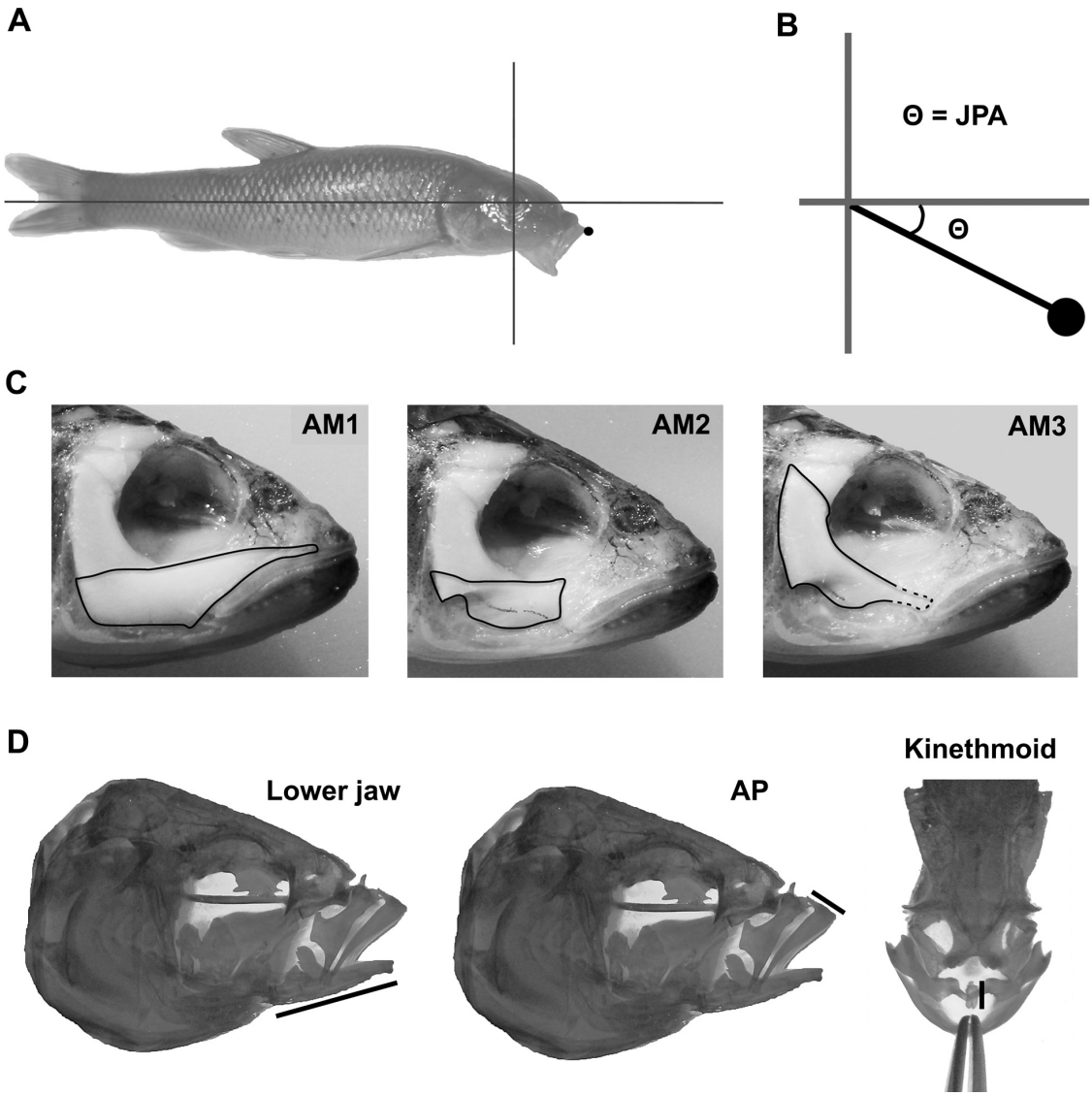

Figure 3.2. Morphological measurements utilized in Chapter III. A) A specimen of *Nocomis micropogon* displaying the positioning of the fish prior to taking the JPA measurement. B) A diagram of the JPA measurement. The black dot marks the tip of the maximally protruded premaxilla for an individual being measured. C) The position of the three adductor mandibula (AM) muscles. AM1 attaches via a thin ligament to the maxilla and functions during cyprinid jaw protrusion. AM2 attaches to the posterior edge of the dentary. AM3 attaches medially on the dentary. D) The three osteological measurements taken. AP denotes the ascending process of the premaxilla.

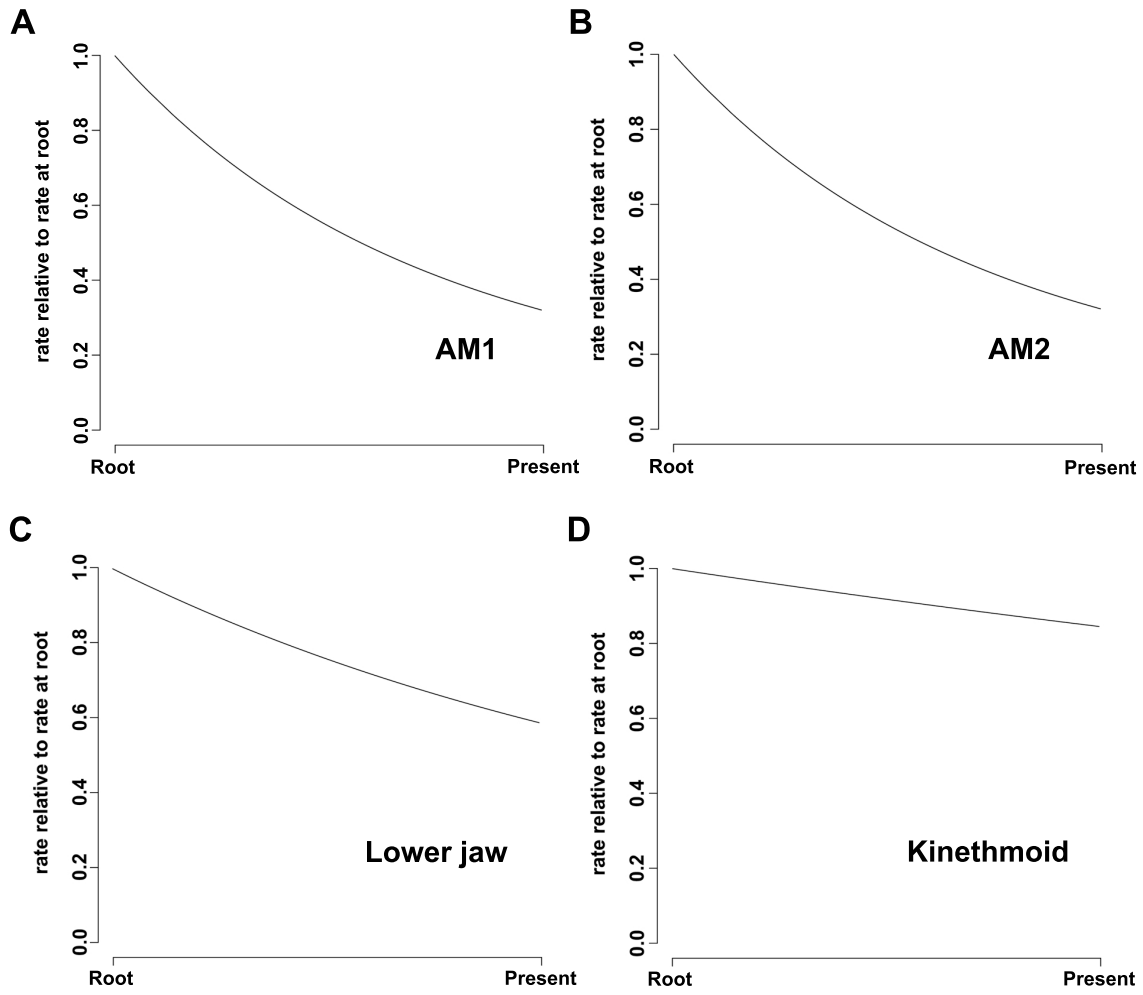

Figure 3.3. Relative evolutionary rate decrease after the pelagic shift and towards the present for four morphological correlates of JPA. A) AM1 mass B) AM2 mass C) lower jaw length D) kinethmoid length. These plots are based on transformations of the Cyt*b* MCC phylogeny using model average estimates of α for each trait (see Chapter III Methods). The "Root" denotes the most recent common ancestor of the clade arising after the benthic to pelagic transition.

VITA

Phillip R. Hollingsworth Jr. was born in Nashville, TN to Phillip Hollingsworth Sr. and Connie Tillis. He attended St. Edward's Catholic school until sixth grade when he moved to Branson, MO where he attended junior high and high school. After graduating from Branson High School, he completed an undergraduate degree in Biology with a minor in Chemistry at Southwest Missouri State University in 2003. Phillip then accepted a graduate teaching assistantship in the Ecology and Evolutionary Biology (EEB) Department at the University of Tennessee, Knoxville (UTK) to work with Dr. David Etnier and Dr. Thomas Near on a M.S. degree. After completing his Master's, he worked as an aquatic biologist for TVA and then a lab technician in the lab of Dr. C. Darrin Hulsey at UTK. Phillip began work on his Ph.D. dissertation in the EEB department at UTK in 2009.