

5-2007

# Geographic and Temporal Diversification Patterns in the Barcheek Darter Species Group

Phillip Ray Hollingsworth Jr.
*University of Tennessee - Knoxville*

To the Graduate Council:

I am submitting herewith a thesis written by Phillip Ray Hollingsworth Jr. entitled "Geographic and Temporal Diversification Patterns in the Barcheek Darter Species Group." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Ecology and Evolutionary Biology.

David A. Etnier, Major Professor

We have read this thesis and recommend its acceptance:

Thomas J. Near, James A. Fordyce

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

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

To the Graduate Council:

I am submitting herewith a thesis written by Phillip Ray Hollingsworth Jr. entitled "Geographic and Temporal Diversification Patterns in the Barcheek Darter Species Group." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Ecology and Evolutionary Biology.

David A. Etnier

Major Professor

Christine R. B. Boake

Department Head

We have read this thesis
and recommend its acceptance:

Thomas J. Near

James A. Fordyce

Accepted for the Council:

Linda Painter

Interim Dean of the Graduate School

(Original signatures are on file with official student records)

# GEOGRAPHIC AND TEMPORAL DIVERSIFICATION PATTERNS IN THE BARCHEEK DARTER SPECIES GROUP

A Thesis
Presented for the
Masters of Science
Degree
The University of Tennessee, Knoxville

Phillip Ray Hollingsworth Jr.
May 2007

## DEDICATION

This thesis is dedicated to my entire family for their nurture and encouragement, and to my late grandfather Robert Henderson Hollingsworth Jr. and our eternal quest for “Big John”.

## ACKNOWLEDGEMENTS

I would like to thank all of those who assisted me in my pursuit of a Master of Science degree. I would especially like to thank Dr. David Etnier and Dr. Tom Near for their invaluable guidance, friendship, and support. I would like to thank Dr. Rex Strange and Dr. Jean Porterfield for donating DNA samples for this study. I would also like to thank my lab mates Ben Keck, Christen Bossu, Richard Harrington, and Jacob Kendrick for their assistance in collecting darters for this project and our time spent around the campfire somewhere between McMinnville and Tampa Bay.

I would finally like to apologize to and thank my family and Sarah Farnsley who now have to live with an ichthyologist.

## ABSTRACT

The purpose of this study was to investigate geographic and temporal diversification patterns in the barcheek darter species group. Specifically, my two questions were "Is there geographical structure of alleles or haplotypes within currently recognized species that is suggestive of unrecognized, or cryptic, species diversity within the clade?" (geographic diversification pattern) and "How old are inter- and intraspecific divergence events in the evolutionary history of the clade?" (temporal diversification pattern). A three gene dataset from 159 barcheek individuals of two mitochondrial coding regions, cyt *b* and ND2, along with a nuclear intron, S7, was analyzed using parsimony and Bayesian phylogentic methods to answer the first question. Divergence times were estimated using fossil calibration of this Bayesian phylogeny in order to answer the second question. Three barcheek species were found to have significant population structure suggestive of cryptic species diversity. *E. basilare* in particular was recovered as being comprised of five reciprocally monophyletic clades endemic to each of the major tributaries to the upper Caney Fork River. Inter- and intraspecific divergence events were found to be relatively old in the clade, nearly all pre-Pleistocene, with a crown node age estimated at 12.68 mya. These results are discussed in light of the present understanding of the tempo of diversification in the darter radiation.

## Table of Contents

## List of Tables

## List of Figures



.

## Chapter I - Introduction

### The barcheek darter species group and the subgenus *Catonotus*

The barcheek darter species group consists of seven described species that are distributed in a mosaic of adjacent, allopatric ranges along the Cumberland, Tennessee, and Green River systems in Tennessee and Kentucky (Figure 1) (Page and Schemske 1978, Etnier and Starnes 1993, Page et al. 2003). All figures and tables in this thesis are presented in the appendix. Species in the group include, *Etheostoma barbouri*, *E. basilare*, *E. derivativum*, *E. obeyense*, *E. smithi*, *E. striatulum*, and *E. virgatum*. The barcheek darters belong to the subgenus *Catonotus* of *Etheostoma*, which also contains the fantail and spottail darter species groups (Bailey and Gosline 1955, Kuehne and Small 1971, Page 1975, Braasch and Mayden 1985).

*Catonotus* is a clade of 20 species characterized by morphological and molecular synapomorphies, as well as a novel breeding behavior that is thought to be highly derived among darters (Page 1985, Porterfield et al. 1999). Territorial males build nests under the edges of flat rocks in small order streams and display to females. Females choose mates and the two invert themselves to lay and fertilize eggs on the underside of the rock, after which the male remains to guard the eggs until hatching (Page and Bart 1989). During the breeding season males of many *Catonotus* species develop white or yellow swellings on the tips of different dorsal fin elements (Page 1983, Etnier and Starnes 1993). These are hypothesized to function as egg mimics to exploit the preference of females to lay eggs in nests that already contain eggs. This female preference has been empirically

demonstrated in aquarium experiments with two *Catonotus* species (Page 1974, Knapp and Sargent 1989).

The barcheek species group is morphologically defined within *Catonotus* by the lack of these egg-mimicking fins. They also differ from other *Catonotus* by having fins with red and blue pigmentation and by having a distinctive bar shaped pattern on their operculum (Figure 2) (Page 1983, Braasch and Mayden 1985, Etnier and Starnes 1993). This bar becomes red and white on males during the breeding season and is hypothesized to function as an independently evolved form of egg-mimicry (Page 2000). Porter et al. (2002) discovered a third form of egg-mimicry that male barcheek darters have evolved in order to exploit this female preference. Using field observations of nesting *E. virgatum* coupled with a microsatellite paternity analysis they found that the number of white spots (egg mimics) on the pectoral fins of breeding males was strongly correlated with the number of offspring he sired (Porter et al. 2002). This evidence suggests that strong sexual selection has likely influenced diversification the barcheek group (Page et al. 1992, Porter et al. 2002).

Another aspect of *Catonotus* ecology that has potentially influenced diversification in the subgenus is the extreme habitat specificity for rocky headwaters exhibited by 18 of the 20 species, including all seven species of the barcheek group (Page et al. 1992). Page et al. (1992) note:

> "Few other daters occupy headwaters, and no group of darters, with the possible exception of the *E. spectabile* group of the subgenus *Oligocephalus*, has specialized in this respect to the degree found in *Catonotus.* Populations of

> headwater fishes tend to be isolated from one another, and the restricted gene flow that results facilitates differentiation. Because of their restriction to patchy habitat, ecologically specialized species such as those of *Catonotus* are expected to show pronounced geographic variation and a propensity to speciate."

An evolutionary history of strong sexual selection coupled with isolation by habitat specificity seems to explain the large number of small range endemic species that make *Catonotus* one of the most speciose subgenera of *Etheostoma*. Aside from *E. flabellare* which ranges over much of eastern North America, all *Catonotus* species display a patchwork of restricted allopatric ranges throughout tributaries to the Cumberland, Tennessee, and lower Ohio Rivers (Etnier and Starnes 1993). Interspecific competition among lineages for habitat space is also hypothesized to play a large role in the origin and maintenance of these allopatric distributions (Page and Schemske 1978).

Understanding diversification patterns in this speciose and curiously distributed clade is an important step in reconstructing the evolutionary history of the highly diverse fish fauna of eastern North America. The purpose of this study was to use molecular phylogenetics and fossil calibrated molecular divergence time estimates to investigate geographic and temporal patterns of diversification in a monophyletic subgroup of *Catonotus*, the barcheek darter clade. Specifically, my two questions were: "Is there geographical structure of alleles or haplotypes within currently recognized species that is suggestive of unrecognized, or cryptic, species diversity within the clade?" (geographic diversification pattern); and "How old are

inter- and intraspecific divergence events in the evolutionary history of the clade?" (temporal diversification pattern). The second section of this introduction discusses the problem of cryptic species as they apply to the eastern North American ichthyofauna and the barcheek darters in particular. The third and final section of the introduction discusses what we know about the timing of diversification in eastern North American fishes and what we may learn by estimating divergence times in the barcheek darter clade.

**The problem of cryptic species as they apply to eastern North American fishes**

Mayr (1963) defined *sibling species* as "morphologically similar or identical populations that are reproductively isolated" and discussed their prevalence among animal populations. Cryptic *species* has recently become the more common term applied to this evolutionary phenomenon in which morphological stasis in a lineage causes species diversity in that lineage to go underestimated until more closely investigated using genetic data or more rigorous morphological methods. Molecular phylogenetic analysis of DNA sequence data has revealed the presence of cryptic species across the metazoan tree of life from North American springsnails (Liu et al. 2003) and Antartic icefish (Bernardi and Goswami 1997) to neotropical harlequin beetle-riding pseudoscorpions (Wilcox et al. 1997). The exceptionally diverse radiations of freshwater fishes in eastern North America are no exception. Molecular and morphological analysis continues to uncover cryptic diversity in several of the most speciose lineages, including minnows (Pera and Armbruster 2006), madtoms (Egge and Simons 2006), and darters (Wood and Raley 2000, Switzer and Wood 2002, Page et al. 2003).

Among the darters, few arguably monophyletic clades exhibit the extreme morphological similarity found between members of each of the three species groups of *Catonotus* (Etnier and Starnes 1993). This morphological stasis caused species diversity in the group to be underestimated for many years, which is reflected in their taxonomic history. Fifteen of the twenty recognized *Catonotus* species have been described since 1971 (Nelson et al. 2004). For example, the spottail darter species group was long considered to be one wide-ranging species, *E. squamiceps*. The 1980s saw the number of described species in the spottail group jump from one to ten based primarily on differences in the second dorsal fin of breeding males (Page et al. 1992). Females in this species group are virtually identical, making it nearly impossible to identify them to species without a working knowledge of the largely allopatric ranges of these 10 cryptic species (Etnier and Starnes 1993).

The barcheek group has also had cryptic species recognition increase the number of species in the group in recent years. Using molecular and morphological data, Page et al. (2003) split the three disjunct populations of *E. virgatum* into three species. The barcheek species in the Rockcastle River and Buck Creek of Kentucky retained the senior *E. virgatum*, while the Caney Fork endemic was described as *E. basilare* and the populations farther down the Cumberland River were named *E. derivativum*. The presence of horizontal brown lines on the sides of these three species had caused earlier workers to overlook species diversity in the barcheek darter clade by lumping these three cryptic species under the name *E. virgatum* (Page et al. 2003).

Recently, genetic data has suggested that there may be additional species diversity currently unrecognized in the barcheek darter clade. In an analysis of AFLP (amplified

fragment length polymorphism) data, Mendelson and Simons (2006) found significant intraspecific population structure throughout the barcheek darter group. These authors sampled multiple populations from all seven barcheek species and found that out of fifteen populations from which they had sampled multiple individuals, twelve received significant statistical support as being monophyletic (Mendelson and Simons 2006). In this study Mendelson and Simons (2006) also presented new hypothesized interspecific relationships within the barcheek group based on AFLP data and discussed why they felt that using AFLP data was superior to DNA sequence data (Page et al. 2003) in inferring the barcheek phylogeny. However, their trees were largely congruent with the mtDNA + nDNA maximum likelihood analysis from Page et al. (2003) although they did offer more statistical support for several relationships. In light of these two studies it was not in my interest to propose new interspecific relationships but instead to use mtDNA and nDNA data to infer intraspecific relationships in the barcheek group. I combined previously published data from Page et al. (2003) for two mtDNA genes and a nDNA intron with data from the same loci for 147 additional barcheek specimens representing multiple populations of all seven species. I performed parsimony-based and Bayesian phylogenetic analyses on this sequence data to search for divergent intraspecific population structure suggestive of cryptic species diversity. Pilot data for this study, as well as AFLP data from Mendelson and Simons (2006) and morphological data from Page et al. (2003), suggests that *E. basilare*, in particular, may be a complex of cryptic species currently recognized under a single name. Therefore, *E. basilare* was thoroughly sampled and 99 individuals were obtained from 13 populations representing the entire

geographic range of this species to be included in phylogenetic analyses. Figure 3 is a detailed map of the range of this species with all collection localities mapped.

**Timing of diversification in eastern North American highland fishes**

Eastern North America is home to the most diverse temperate freshwater fish fauna in the world (Briggs 1986, Lundberg et al. 2000). The darters (Percidae: Etheostomatini) make up a large portion of this diversity as they have diversified into impressive evolutionary radiation made up of more than 220 species, all of which are endemic to eastern North America (Lundberg et al. 2000). The majority of darter diversity is concentrated in the streams and rivers draining three disjunct areas of highland terrain. These Central Highlands as defined by Mayden (1985) and Wiley and Mayden (1985) are comprised of the Interior Highlands (Ozark and Ouachita Mountains) in the west and the Eastern Highlands (southern Appalachian Mountains and associated plateaus) in the east.

The temporal evolution of this highland darter fauna is poorly understood. However, a few hypotheses have been presented regarding the chronology of the darter radiation. Although darters first appear in the fossil record in the late Pleistocene (Smith 1981, Cavender 1986), the majority of these aforementioned hypotheses place the origins of diversity in the clade prior to the Pleistocene either explicitly (Page 1983) or indirectly (Pflieger 1971, Mayden 1985, Wiley and Mayden 1985, Mayden 1987a, Mayden 1987b, Mayden 1988, Strange and Burr 1997, Near et al. 2001, Near and Keck 2005). Page (1983) states explicitly “Darters, unknown from the [pre-Pleistocene] fossil record, . . . have originated and diversified since the Pliocene.” However, he neither gives nor cites any evidence for this hypothesis. The indirect estimates of the timing of darter

diversification come from a paradigm of North American ichthyology known as the Central Highlands Vicariance Hypothesis (CHVH) (summarized in Mayden 1988), the opposing Pleistocene Dispersal Hypothesis (reviewed in Mayden 1987b), and recent tests of these two biogeographical theories (Strange and Burr 1997, Near et al. 2001, Berendzen et al. 2003, Near and Keck 2005). These hypotheses indirectly and ambiguously date the origination of the majority of darter diversity to before the Pleistocene for the following reasons: 1) The CHVH proposes that a widespread and already diverse fish fauna, including darters, inhabited a large, interconnected river system draining a contiguous Central Highlands prior to the Pleistocene. According to the CHVH, glacial cycles during the Pleistocene fragmented the Central Highlands and its river system into the disjunct highland areas and river systems of today leading to speciation in many fish lineages as sister species in the highlands became isolated by unsuitable, lowland habitat between highland regions (Mayden 1988). 2) The Pleistocene Dispersal Hypothesis suggests the Eastern Highlands were the center of origin for a diverse ancestral fish fauna. Glacial advance presumably allowed certain Eastern Highland fish lineages to disperse into the Interior Highlands by reducing flow in the river systems draining the area between the disjunct highland regions and creating dispersal pathways across continuous high-gradient watersheds. Allopatric speciation occurred when these dispersal pathways were destroyed by the increased flows from the glacial retreat, leaving lineages isolated in the disjunct highlands (reviewed in Mayden 1987b).

Recently, molecular phylogenetic studies of fishes distributed in the Central Highlands have uncovered evidence that suggests Pleistocene vicariance, as well as

recent dispersal, is necessary to explain the contemporary composition of these fish communities in the Interior and Eastern Highlands (Strange and Burr 1997, Near et al. 2001, Berendzen et al. 2003, Near and Keck 2005). So while both of these hypotheses appear valid in explaining the distribution and ages of select Central Highland fish lineages, neither attempts to estimate the age of the ancestral diversity that they both assume was present in the Central Highlands prior to the Pleistocene. Regardless of whether the Eastern Highlands served as a center of origin for Central Highland fish clades which later dispersed into the Interior Highlands during the Pleistocene, or whether the diversity was already present in the Interior Highlands and underwent vicariance during the Pleistocene, the question becomes: "What is the age of darter lineages endemic to unglaciated and climatically stable regions of the Central Highlands?"

Two recent studies using fossil calibrated molecular divergence time estimates of different darter clades suggest some darter lineages are quite old (Near and Keck 2005) and some exhibit younger ages (Near and Bernard 2004). Near and Keck (2005) estimated that diversification of the darter clade *Nothonotus* began approximately 18.5 million years ago (mya) while testing temporal predictions of the CHVH with this speciose clade. In a different study, Near and Bernard (2004) found that the logperch darter clade diversified from a recent common ancestor into 10 species in just over 4 million years through rapid allopatric speciation.

The first goal of this thesis was to estimate divergence times between species and populations in a third clade of darters, the barcheek species group, using fossil calibration methods similar to those of Near and Bernard (2004) and Near and Keck (2005). The

majority of barcheek darter species diversity is concentrated in the Cumberland River (Etnier and Starnes 1993) (Figure 1), which is one of very few rivers in eastern North America whose configuration is thought to have remained relatively unaltered through the Pleistocene glacial cycles (Mayden 1988). Figure 4 (Mayden 1988) displays the hypothesized pre-glacial drainage configuration of rivers in eastern North America and the Cumberland River is shown running much as it does now complete with major tributaries such as the Caney Fork River, Big South Fork, and Rockcastle River all of which are inhabited by barcheek darters. The portions of the two systems outside of the Cumberland River in which barcheeks occur, the middle Duck River (*E. striatulum*) and the upper Green River system (*E. barbouri*), are also shown with very similar pre-glacial configurations in Figure 4. Like the darter radiation as a whole, few statements have been made regarding the ages of divergence events within *Catonotus* and the barcheek darter species group. Braasch and Mayden (1985) stated that diversity within the subgenus *Catonotus* " . . . may be quite old (mid-Tertiary)". These authors go on to discuss the historical biogeography of the barcheek darters and acknowledge the difficultly in accurately aging diversification in this clade without evidence of significant geologic events in the history of the relatively ancient and stable Cumberland River system (Braasch and Mayden 1985). They do suggest however that stream capture during the Pleistocene must be responsible for the birth of the species *E. striatulum* in the Duck River system of the Nashville Basin (Braasch and Mayden 1985), while Page and Braasch (1976) when describing *E. smithi* as a new species distinct from *E. obeyense* suggested that these species " . . . presumably have differentiated in relatively recent time." Given the stability of the Cumberland River system and the absence of geologic

events (pre-glacial or glacial) as potential biogeographical proxies for divergence times in Cumberland River endemics it becomes interesting to use a fossil calibrated molecular phylogeny to estimate absolute divergence times in the barcheek species group. These results should give us a better understanding of how long darter lineages have occupied this stable region of eastern North America.

## Chapter II - Methods

### Specimen collection and DNA sequencing

The author, along with the help of Drs. Thomas Near, Rex Strange, Jean Porterfield, and several others, collected 147 individuals representing all seven recognized barcheek darter species using standard seining techniques. The common fish anesthetic MS-222 was used to sedate fish prior to tissue acquisition. Tissue samples were obtained either by preserving a whole fish in absolute ethanol or removing an individual's right pectoral fin and storing it in absolute ethanol in a 2 mL micro-centrifuge tube. If a fin tissue sample was taken from an individual the remaining specimen was kept as a voucher. Tissue samples were kept at 4°C for long-term storage, and voucher specimens were deposited into the University of Tennessee Research Collection of Fishes (UT), the Illinois Natural History Survey (INHS), or the North Carolina State Museum (NCSM). Collection localities are mapped in Figure 1 and are presented along with museum voucher information (when available) in Table 1. DNA isolation was performed on the 147 tissue samples using standard phenol-chloroform extraction and ethanol precipitation procedures along with Qiagen DNAeasy tissue kits. Purified genomic DNA was stored in 1X TE buffer and kept at -20°C for long-term storage. Complete coding regions of the mitochondrial NADH dehydrogenase 2 (ND2) and cytochrome *b* (cyt*b*) genes were PCR amplified using primers and conditions given in Kocher et al. (1995) and Near et al. (2000), respectively. The primers and conditions given in Kocher et al. (1995) also amplified the tRNA regions flanking the ND2 gene. The first intron of the S7 ribosomal protein (S7) was PCR amplified using primers and conditions from Chow and Hazama (1998) in order to include a nuclear marker in my

analyses. The 2924 total nucleotides sequenced for this study can be decomposed as: 1140 cyt*b*, 1047 ND2, 210 tRNA, and 527 S7. Previously published data for these same gene regions for 12 additional barcheek individuals was downloaded from GenBank. Locality information and museum voucher information for these previously published sequences are given in Table 1 and GenBank accession numbers for these sequences are given in Table 2.

Based on previous phylogenetic studies of the subgenus *Catonotus* (Page 1975, Braasch and Mayden 1985, Porterfield et al. 1999, Page et al. 2003, Mendelson and Simons 2006), I chose to root the hypothesized barcheek trees with five species representing the two other species groups in the clade: spottail and fantail darters (Etnier and Starnes 1993). Data from the loci used in this study was downloaded from GenBank for these other 5 *Catonotus* species (Page et al. 2003), and accession numbers for these sequences are presented in Table 2. These five outgroup species were each represented by one individual in all analyses. Lastly, ND2 and S7 sequence data for 46 individuals representing all 32 contemporary species of the freshwater fish family Centrarchidae was downloaded from GenBank (Near et al. 2004) for use in molecular divergence time (MDT) analysis of the barcheek darter clade (see below). GenBank accession numbers for these centrarchid sequences are given in Table 2.

**Parsimony-based phylogenetic analyses**

The program ClustalX (Thompson et al. 1997) was used to align all sequence data. The alignment between *Catonotus* species was straightforward and unambiguous. The spottail darters *E. oophylax* and *E. squamiceps* were designated as outgroups in all

phylogenetic analyses. The fantail darters *E. flabellare*, *E. kennicotti*, and *E. percnurum* were included in the ingroup in all analyses as previous results (Page et al. 2003) have suggested paraphyly of the barcheeks with respect to the fantail darters.

Insertions/deletions, heterozygous polymorphic sites, and missing data were ignored in all phylogenetic analyses. The only missing data are approximately 20-30 base pairs from the beginning of the cyt*b* coding sequence for three *E. striatulum* and three *E. basilare* individuals. The nDNA data and mtDNA data (cyt*b*, ND2, and tRNA) partitions were analyzed separately before the two partitions were concatenated (mtDNA + nDNA) and analyzed as a single data set. Nuclear S7 data was analyzed by itself first in order to provide an independent hypothesis of barcheek intraspecific relationships using a non-cytoplasmic marker (Hare 2001; Zhang and Hewitt 2003). I identified all unique S7 alleles, ignoring heterozygous polymorphisms, using TCS (Clement et al. 2000) and subjected this data to a parsimony based heuristic-tree search in PAUP* with TBR branch swapping and 100 random addition sequence replicates. The MulTrees option, which saves multiple optimal trees from each addition sequence replicate, was turned off to expedite computation time due to the large number of taxa. Given the low phylogenetic resolution offered by the S7 gene (see Results) a 50% majority-rule consensus of all most parsimonious trees was found and support for this tree was assessed by non-parametric bootstrap analysis (Felsenstein 1985) with 1000 "fast" stepwise addition pseudoreplicates. Nodes receiving a bootstrap score below 50 were collapsed when presenting the results. Parsimony analysis of the mtDNA (cyt*b*, ND2, and tRNA) and mtDNA + nDNA data sets followed the exact same procedure.

Owing to the higher phylogenetic resolution when analyzing these larger data sets a strict consensus tree of all most parsimonious trees was found for each instead of the less resolute 50% majority-rule consensus. Once again, nodes receiving a bootstrap of less than 50 were collapsed in the presented trees. The nDNA data added little phylogenetic signal when combined with the mtDNA data such that results from these analyses (mtDNA + nDNA and mtDNA only) were virtually identical, therefore only the mtDNA + nDNA results are presented. Lastly, TCS was used to create a statistical parsimony based haplotype network of the combined mtDNA (cyt*b*, ND2, and tRNA) data for 99 *E. basilare* individuals.

**Bayesian phylogenetic analysis, tests of nucleotide substitution rate heterogeneity, and molecular divergence time (MDT) estimates**

I employed the method of Near and Bernard (2004) and Near and Keck (2005) in order to estimate divergence times in the barcheek species group. This involved generating a data set that contained both barcheeks and individuals from the freshwater fish clade Centrarchidae. Fossil age information from the centrarchids was then used to date nodes in this phylogeny, allowing absolute divergence time estimates in the barcheek clade to be inferred from Bayesian likelihood branch lengths. The clade Centrarchidae is characterized by two attributes making this analysis possible. The family has been the focus of recent phylogenetic studies yielding large amounts of comparative genetic data (Roe et al. 2002, Near et al. 2003, 2004, 2005, Harris et al. 2005) and they are well represented in the late Cenozoic fossil record (Smith 1981, Cavender 1986). Unfortunately, the darter fossil record is sparse with the earliest verified

fossils appearing in Pleistocene-aged deposits (Smith 1981, Cavender 1986), thereby limiting the utility of darter fossils as calibration points for divergence time estimates. As in Near and Bernard (2004) and Near and Keck (2005), I assumed that using the centrarchid fossils as external calibrations for divergence times in the barcheek darter clade would lead to accurate divergence time estimates due to the relatively close phylogenetic affinity between the darters and Centrarchidae along with my methods for smoothing rate variation across lineages (see below).

In order to obtain a phylogenetic tree upon which to base divergence time estimates I performed the following analyses. The ND2 and S7 data from 159 barcheek individuals and 5 non-barcheek *Catonotus* individuals was combined with data from the same regions for 46 individuals representing all 32 centrarchid species (Near et al. 2004). This data was initially aligned using ClustalX.1 (Thompson 1997) and subsequently MacClade 4.08 (Maddison and Maddison 2002) was used to clean up the alignment by minimizing inferred insertions/deletions. The barcheek cyt*b* and tRNA data were excluded from this analysis as data from the same regions were not available for the same set of centrarchid specimens. In order to account for site-specific rate heterogeneity the ND2 + S7 concatenated data set was separated into four data partitions: 1st codon position of ND2, 2nd codon position of ND2, 3rd codon position of ND2, and S7. These data partitions were each subjected to hierarchical likelihood ratio tests in ModelTest (Posada and Crandall 1998) in order to choose the least parameter-rich model of molecular evolution making the data for each partition the most probable, or likely. Once models were selected for each data partition, the data set was subjected to a partition mixed model (pMM) Bayesian analysis using MrBayes 3.0 (Ronquist and Huelsenbeck

2003). Models chosen in ModelTest were assigned to their appropriate partition using the APPLYTO command in MrBayes. Model parameters for each partition were estimated independently by using the UNLINK and PRSET rates = variable commands in MrBayes. The pMM Bayes search consisted of four chains, three hot and one cold, and ran for 5 X $10^6$ generations with trees being sampled and saved every 100 generations. Posterior probabilities of nodes in the pMM Bayesian tree were estimated using the metropolis-coupled Markov chain Monte Carlo (MC3) algorithm with the frequency of occurrence of a particular node in all trees sampled after the "burn-in" at $1X10^6$ generations representing its posterior probability (Larget and Simon 1999, Ronquist and Huelsenbeck 2003).

The ND2 and S7 concatenated data set was tested for significant rate heterogeneity of molecular evolution across constituent *Catonotus* and centrarchid lineages in order to determine whether it would be necessary to account for this variable in estimating divergence times. This was achieved by first choosing an optimal model of molecular evolution for the concatenated ND2 and S7 dataset using ModelTest. Then the consensus post burn-in pMM Bayes tree was obtained using the SUMT command in MrBayes and imported into PAUP* so that model parameters for the analysis could be estimated. Using the chosen model and estimated parameters the likelihood values of the pMM Bayes tree with the molecular clock enforced versus not enforced were compared using a LRT with a chi squared distribution and s-2 degrees of freedom, where s equals the number of taxa in the analysis. Once I had obtained a tree and determined there to be significant rate heterogeneity across it I utilized the penalized likelihood method of Sanderson (2002), as implemented through the program r8s, to account for this rate

heterogeneity when estimating divergence times (Sanderson 2003, Near and Sanderson 2004). Penalized likelihood corrects for autocorrelation of rate transformation between ancestor and descendent branches in a phylogeny by imposing a smoothing parameter that allows molecular evolution rates to vary across the tree without this variation becoming too extreme (Sanderson 2002). This smoothing parameter is chosen by a cross-validation procedure in r8s (Sanderson 2003). Using the SUMT command again in MrBayes the strict consensus post-burn in tree and likelihood branch lengths were imported into r8s. Six centrarchid fossils were then used as calibration points in r8s to estimate absolute divergence times on the barcheek + centrarchid tree. A novel cross validation analysis in Near et al. (2005) suggested these six to be the most consistent with one another out of ten fossils analyzed when dating the centrarchid phylogeny. See Near et al. (2005) for exact placement of these fossils in the centrarchid phylogeny. Fossil ages and sources are given in Table 3. The program, r8s, allows these calibration points to be set as fixed ages or bracketed by minimal and/or maximal age constraints. I set three fossil calibration points as fixed and three as minimal age constraints (Table 3).

## Chapter III - Results

### nDNA parsimony-based phylogenetic analyses

The computer program TCS identified 37 unique S7 alleles, ignoring heterozygous polymorphisms, from the 159 barcheek individuals sampled. Only 67 of 529 nucleotide positions from this locus were found to be parsimony-informative when analyzed with S7 data from the 5 *Catonotus* "rooting" individuals. Parsimony analysis of these characters yielded 67 most parsimonious trees that were 157 steps (base pair changes) long. A 50% majority rule consensus tree was generated from these 67 trees and is presented in Figure 5. Bootstrap support values were all below 75 and are not presented. Frequency of occurrence of each node on the 67 most parsimonious trees is given as a less statistically rigorous measure of support in Figure 4. The barcheek species group was recovered as monophyletic in this analysis as they were in Page et al. (2003). However, only three of seven species of barcheeks were recovered as monophyletic: *E. barbouri*, *E. basilare*, and *E. striatulum*. This intraspecific paraphyly based on S7 data was not observed in Page et al. (2003), most likely due to the fact that that study included data from no more than two individuals per species. While parsimony analysis of the nDNA data supports a monophyletic barcheek darter species group it offers little resolution for inter- and intraspecific relationships within this clade when multiple populations from each species are included in phylogenetic analysis. The only intraspecific population structure, albeit weakly supported, found among S7 allels was within *E. basilare*. This parsimony analysis recovered a clade consisting of all *E. basilare* individuals sampled from the Caney Fork River proper as sister to a clade consisting of all other *E. basilare* sampled.

**mtDNA + nDNA parsimony-based phylogenetic analyses**

The concatenated mtDNA + nDNA data set consisted of 917 out of 2926 parsimony informative characters. Parsimony analysis of this dataset produced 23 most parsimonious trees. These most parsimonious trees were 3158 steps long. The strict consensus of these 23 most parsimonious trees is given along with bootstrap support values for nodes in Figure 6. Well-supported intraspecific population structure was recovered within three barcheek darter species by this parsimony analysis, *E. obeyense*, *E. smithi*, and *E. basilare*. These are three of the four barcheek species in which more than 4 individuals and two populations were sampled for this study. Figure 7 is a phylogram of one of the 23 most parsimonious trees displaying the relatively long intraspecific parsimony branch lengths in these three species. *E. smithi* was recovered as two monophyletic clades receiving bootstrap scores of 100. Individuals sampled from Spring Creek were recovered as a monophyletic clade sister to a clade consisting of *E. striatulum* and the rest of the *E. smithi* populations sampled. Bootstrap support for this paraphyly of *E. smithi* was low at 62.

Within *E. obeyense* all three populations from which multiple individuals were sampled, West Fork of Obey River, Mill Creek, and the Little South Fork River, were recovered as monophyletic with bootstrap support > 97.

Five clades all receiving bootstrap scores of 100 were recovered within *E. basilare*. These clades are each confined to major tributaries of the Caney Fork system (Figure 3). Individuals sampled from the Barren Fork River headwaters in Cannon and western Warren counties are the most basal lineage according to this hypothesis with the next divergence taking place between all individuals sampled from the Collins River and

those sampled from the Caney Fork proper above its confluence with the Collins River. Genetic subdivision is also present within the upper Caney Fork River proper where three reciprocally monophyletic groups correspond to three major tributaries, Calfikiller River, Rocky River, and Cane Creek plus extreme upper Caney Fork River. After these results I calculated uncorrected parsimony distance at cyt *b* between the two most divergent individuals of *E. basilare* (Duke Creek and Scott Creek). The uncorrected parsimony distance of 8.5% between these two individuals is among the highest intraspecific divergence in cyt *b* reported in darters, and is certainly the highest on this small of a geographic scale (Wood and Raley 2000, Kinzinger et al 2001, Near et al. 2001, Switzer and Wood 2002).

Only four nodes were poorly resolved in this combined data parsimony phylogeny and the topology of the tree representing interspecific relationships is largely the same as the same analysis in Page et al (2003). One poorly resolved node was that concerning the monophyly of the barcheeks. As in Page et al. (2003), the barcheeks were recovered as paraphyletic with respect to the fantail darters in this parsimony analysis. *E. barbouri* was recovered as sister to a collapsed polytomy containing the fantail darters, *E. basilare*, and a clade of all other barcheek species. However, this node uniting the fantails and barcheeks minus *E. barbouri* received low bootstrap support with a score of 66. I agree with Page et al. (2003) and Mendelson and Simons (2006) in presuming this is due to homoplasy and in accepting the barcheeks as monophyletic. Two other nodes received low boostrap support. The node representing the most recent common ancestor (MRCA) of *E. obeyenese*, *E. derivativum*, *E. smithi*, and *E. striatulum* scored 65 while the node rendering *E. smithi* paraphyletic with respect to *E. striatulum* scored 62.

The population structure observed in *E. basilare* warranted closer inspection. The computer program TCS was used to convert sequence data for the combined mtDNA regions (cyt*b*, tRNA, and ND2) into haplotype networks for each of the five reciprocally monophyletic populations recovered from within *E. basilare* in the parsimony analysis. These haplotype networks are presented in Figure 8.

**Bayesian phylogenetic analysis, tests of nucleotide substitution rate heterogeneity, and molecular divergence time (MDT) estimates**

The first step in generating a Bayesian phylogenetic tree upon which to base divergence time estimates was to choose the appropriate model of sequence evolution for the four data partitions (ND2 $1^{st}$, $2^{nd}$, and $3^{rd}$ codon,1 and S7) used in this portion of the study. LRTs in ModelTest chose the same general time reversible (GTR) model with the added parameters of proportion of invariant sites and gamma distributed substitution rates for the three ND2 partitions (Table 4). The model chosen for the S7 partition was TrN plus gamma distributed substitution rates (Table 4). The pMM Bayesian analysis was then ran and the post burn-in consensus tree was tested for significant rate heterogeneity of molecular evolution across lineages. The likelihood score of this tree using the GTR + I + G model chosen for the concatenated ND2 + S7 data set (Table 4) was -24556.34 with the molecular clock constraint enforced. The likelihood score of the tree without the molecular clock enforced was -23997.47 ($\chi^2$ = 1117.76, d.f. 208 p <<< 0.005), indicating significant rate heterogeneity of molecular evolution across centrarchid and barcheek lineages. The program r8s was then used to estimate divergence times in the barcheek clade while accounting for this rate heterogeneity. Figure 9 is a chronogram of the pMM

Bayesian tree dated using r8s and the centrarchid fossil calibration points. The centrarchid portion of the tree has been removed as it is not relevant to this study.

Two major differences in topology are present between this Bayesian tree and the maximum parsimony tree generated earlier in this study. The barcheeks were resolved as monophyletic with *E. barbouri* and *E. basilare* forming the sister group to all other barcheek darters. Also, *E. smithi* was recovered as monophyletic in this analysis. All but two nodes in the barcheek portion of the tree received posterior probability scores of 100 from MC3 sampling. The node uniting *E. barbouri* and *E. basilare* as sister lineages received a posterior probability of 95 indicating significant support. The node representing the MRCA of *E. obeyense*, *E. derivativum*, *E. smithi*, and *E. striatulum* received a posterior probability score of 87 indicating a lack of support for this node. Interestingly, this node received a low bootstrap score (67) in the parsimony analysis of the total data set.

Divergence time estimates are given in Table 5. Divergence time analysis suggests the age of the MRCA for all barcheek darters to have lived at 12.68 mya. Interspecific divergences within the clade range from 11.87 mya between *E. basilare* and *E. barbouri* to 3.09 between *E. striatulum* and *E. smithi*. Intraspecific divergences within *E. basilare* range from 5.61 when the Barren Fork River populations diverged from populations in the rest of the Caney Fork River system. The Collins River populations diverged from the Caney Fork River populations at 3.70 mya, the Calfkiller River population from the Rocky River and Cane Creek populations at 2.47 mya, and finally the Rocky River populations from the Cane Creek and upper Caney Fork populations at 1.49 mya. Deeply divergent population structure was also uncovered in *E. obeyense* and

*E. smithi*. Although exact age estimates for intraspecific divergences in these species were not obtained; the MRCA for populations comprising each of these species was dated at 4.56 and 2.57 mya respectively.

## Chapter IV - Discussion

### Cryptic species diversity within the barcheek darter group

The first question posed in this thesis was, "Is there geographical structure of alleles or haplotypes within currently recognized species that is suggestive of unrecognized, or cryptic, species diversity within the barcheek darter clade?" I used a multi-gene phylogeny with multiple populations sampled for all seven barcheek species to answer this question. Parsimony analysis of the nuclear S7 alone data offered little resolution for both inter- and intraspecific relationships within the clade. However, the combined nDNA + mtDNA parsimony analysis revealed significant population structure in several barcheek species. Collecting efforts for this study were focused on the Caney Fork River endemic barcheek species, *E. basilare*, and individuals were obtained representing the vast majority of the range of this species. Five reciprocally monophyletic clades were recovered within *E. basilare*, each with strong statistical support. Each of the five clades is endemic to a major tributary to the system: Barren Fork River, Collins River, Rocky River, Calfkiller River, and the extreme upper Caney Fork River plus Cane Creek. In addition to being monophyletic, parsimony branch lengths were relatively long between these clades in this analysis of combined data. In fact, the 8.5% uncorrected divergence in cyt *b* between the two most divergent *E. basilare* individuals is among the highest reported to date between conspecific darters for this commonly used molecular marker (Wood and Raley 2000, Kinzinger et al. 2001, Near et al. 2001, Switzer and Wood 2002).

Haplotype networks of these populations are generally "star" shaped with several individuals sharing a single haplotype and all other haplotypes being only a few base pair

changes away from the former (Figure 8). Long interpopulation branch lengths and limited molecular diversity within populations is consistent with a long history of reproductive isolation between populations and frequent bottlenecks within populations. Barcheek darters are often collected in small numbers and can be especially patchy in their distribution within their known ranges (Etnier and Starnes 1993, personal obs.). Organisms with small population sizes are more prone to the effects of genetic bottlenecks and this may reflect in haplotype networks for the five reciprocally monophyletic populations of *E. basilare*. The monophyletic and divergent nature of these populations suggests *E. basilare* is comprised of five reproductively isolated populations, or cryptic species, and each should minimally be considered an evolutionary significant unit (ESU) (Moritz 1994).

Formal species recognition within *E. basilare* may be warranted if morphological divergence accompanied this molecular divergence. Page et al. (2003) analyzed morphological data and found slight differences in several meristic counts between individuals pooled from the Caney Fork River proper and those pooled from the Collins River and its tributaries, including the Barren Fork River. According to this analysis the Barren Fork River population is basal to a clade consisting of Collins River populations as sister to Caney Fork River proper populations (Figure 6). Recognition of two species based on currently available morphological data (Page et al. 2003), one from the Collins River and Barren Fork River and one from the Caney Fork River proper, would create a paraphyletic Caney Fork River proper species and a polyphyletic Collins River (plus Barren Fork River) species. A more in depth morphological investigation into this

species may or may not find differences in morphological characters consistent with the phylogeny proposed in this study.

An alternative hypothesis to the reproductive isolation between populations suggested by this analysis is that female philopatry has caused mitochondria to get captured in these tributaries leading to apparent differentiation of populations that are not truly reproductively isolated (Hoelzer 1997). This analysis was heavily based on mtDNA data and this is possible. However, Mendelson and Simons (2006) found similar population structure within *E. basilare* based on AFLP data that is largely non-cytoplasmic, and should be resistant to difficulties in interpretation due to sex-biased gene flow. Their AFLP data recovered three reciprocally monophyletic populations that correspond to three clades also recovered in this study, namely, a Collins River clade, a Calfkiller River clade, and a Barren Fork River clade (Mendelson and Simons 2006). Their results concerning interpopulation relationships within *E. basilare* indicated that the Collins River and Barren Fork River clades were more closely related to one another with the Calfkiller River clade being basal to these two. Though the topology of the tree representing phylogenetic relationships between these monophyletic populations of *E. basilare* differs between this study and that of Mendelson and Simons (2006), the recovery of these reciprocally monophyletic populations from two very different types of genetic data analysis is strong evidence in support cryptic species diversity within the species *E. basilare*.

Significant population structure was also found in *E. obeyenese* and *E. smithi* which were two of the better geographically sampled species in this study. These results coupled with those from Mendelson and Simons (2006) suggests cryptic species diversity

is prevalent in the barcheek darter clade. Monophyletic clades restricted to single tributary systems were found in three barcheek species. The barcheek species group as currently recognized is comprised of seven species with small, adjacent, and allopatric ranges. The evolutionary process that hypothetically caused, and are currently maintaining, this allopatric diversity seem to be taking place on a smaller geographic scale than previously recognized, leading to cryptic and "micro-endemic" barcheek darter species.

**Timing of diversification in the barcheek darter species group**

The second question posed in this thesis was, "How old are inter- and intraspecific divergence events in the evolutionary history of the clade?". The endemism of the barcheeks in one of the most biogeographically stable regions of eastern North America (Cumberland River, middle Duck River, and upper Green River systems) makes using a fossil calibrated molecular phylogeny to estimate divergence times in this group interesting for two reasons. First, these river systems lack sufficient geological evidence to make accurate statements regarding the biogeography and speciation of their fish lineages without using fossil calibrated molecular phylogenies. Second, with regards to the Central Highlands Vicariance Hypothesis, estimating divergence times in lineages endemic to stable river systems in the Eastern Highlands should give us insight into how long before the Pleistocene certain elements of the contemporary darter fauna diversified. In light of recent studies using similar methods to estimate divergence times in other darter clades (Near and Bernard 2004, Near an Keck 2005), estimating divergence times in the barcheeks is another step in reconstructing the temporal development of diversity in the radiation of darters across eastern North America. Diversification in the barcheek

darter clade began approximately 12.68 mya with all subsequent speciation events predating the Pleistocene. Also, the majority of the intraspecific genetic structure that is suggestive of cryptic species diversity pre-dates the Pleistocene. Within *E. basilare* diversification began approximately 5.61 mya with the divergence of the Barren Fork River clade from the remaining barcheek darters in the Caney Fork River. Crown node ages for the other two species showing significant genetic structure of populations, *E. obeyense* and *E. smithi*, were 4.56 mya and 2.67 mya, respectively. Species diversity, recognized and unrecognized, is relatively old within the barcheek darter species group.

Similar to *Nothonotus* darters, in which diversification began 18.46 mya (Near and Keck 2005), diversification in the barcheek species group was largely complete by the presumed onset of glaciation 1.8 mya. In the logperch darter species group, the other darter clade for which divergence times were estimated using similar methods, diversification from a common ancestor into 10 species has taken place in less than 5 million years. These three species groups are but a small portion of darter diversity and tell two different stories regarding the temporal development of contemporary darter diversity. Clades such as *Nothonotus* and the barcheeks push the age of the most recent common ancestor of all darters back in geologic time prior to the Miocene, with much of the diversity in these clades present prior to the Pleistocene. The logperch radiation, on the other hand, is more recent and may have been driven by sea level fluctuations during the Pleistocene (Near and Bernard 2004). Future studies using similar methods to estimate divergence times in darters as well as other clades of North American freshwater fish are integral in accurately reconstructing the complex historical development of eastern North America's exceptionally diverse freshwater fish communities.

## LITERATURE CITED

## Literature Cited


Baldwin, B. G., and M. J. Sanderson. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proceedings of the National Academy of Science of the USA 95:9402-9406.

Bailey, R. M., and W. A. Gosline. 1955. Variation and systematic significance of vertebral counts in the American fishes of the family Percidae. University of Michigan Museum of Zoology Miscellaneous Publiclations 93:1-44.

Berendzen, P. B., A. M. Simons, and R. M. Wood. 2003. Phylogeography of the northern hogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. Journal of Biogeography 30:1139-1152.

Bernardi, G., and U. Goswami. 1997. Molecular evidence for cryptic species among the Antartic fish *Trematomus bernacchii* and *Trematomus hansoni*. Antartic Science 9:381-385.

Briggs, J. C. 1986. Introduction to the zoogeography of North American fishes. Pp. 1-16 *in* C. H. Hocutt and E. O. Wiley, eds. The zoogeography of North American freshwater fishes. John Wiley & Sons, New York.

Braasch, M. E., and R. L. Mayden. 1985. Review of the subgenus *Catonotus* with descriptions of two new species of the *Etheostoma squamiceps* species group. Occasional Papers of the Museum of Natural History of the University of Kansas 119:1-83

Cavender, T. M. 1986. Review of the fossil history of North American freshwater fishes. Pp. 699-724 *in* C. H. Hocutt and E. O. Wiley, eds. The zoogeography of North American freshwater fishes. John Wiley & Sons, New York.

Clement, M., D. Posada, and K. A. Crandall. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1660.

Chow, S. and K. Hazama. 1998. Universal PCR primers for S7 ribosomal protein gene introns in fish. Molecular Ecology 7:1247-1263.

Egge, J. J. D., and A. M. Simons. 2006. The challenge of truly cryptic diversity: diagnoses and description of a new madtom catfish (Ictaluridae: *Noturus*). Zoologica Scripta 35:581:595.

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

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791.

Felsenstein, J. 1993. Phylogenetic inference package. University of Washington, Seattle WA.

Golenberg, E. M., D. E. Giannasi, M. T. Clegg, C. J. Smiley, M. Durbin, D. Henderson, and G. Zurawski. 1990. Chloroplast DNA sequence from a Miocene *Magnolia* species. Nature 344:656-658.

Gradstein, F. M., J. G. Ogg, A. G. Smith, F. P. Agterberg, W. Bleeker, R. A. Cooper, V. Davydov, P. Gibbard, L. A. Hinnoy, M. R. House, L. Lourens, H. P. Luterbacher, J. McArthur, M. J. Melchin, L. J. Robb, J. Shergold, M. Villeneuve, B. R. Wardlaw, J. Ali, H. Brinkhuis, F. J. Hilgen, J. Hooker, R. J. Howarth, A. H. Knoll, J. Laskar, S. Monechi, K. A. Plumb, J. Powell, I. Raffi, U. Rohl, P. Sadler, A. Sanfilippo, B. Schmitz, N. J. Shackleton, G. A. Shields, H. Strauss, J. Van Dam, T. van Kolfschoten, J. Veizer, and D. Wilson. 2004. A geologic time scale 2004. Cambridge University Press, Cambrige, UK.

Hare, M. P. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology and Evolution 16:700-706.

Harris, P. M., K. J. Roe, and R. L. Mayden. 2005. A mitochondrial DNA perspective on the molecular systematics of the sunfish genus *Lepomis* (Actinopterygii: Centrarchidae). Copeia 2005:340-346.

Hibbard, C. W. 1936. Two new sunfish of the family Centrarchidae from middle Pliocene Kansas. University of Kansas Science Bulletin 24:177-185.

Hoelzer, G. A. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear gene trees revisted. Evolution 51:622-626.

Kinziger, A. P., R. M. Wood, and S. A. Welsh. 2001. Systematics of *Etheostoma tippecanoe* and *Etheostoma denoncourti* (Perciformes: Percidae). Copeia 2001:235-239.

Knapp, R. A., and R. C. Sargent. 1989. Egg-mimicry as a mating strategy in the fantail darter, *Etheostoma flabellare*: females prefer males with eggs. Behavioral Ecology and Sociobiology 25:321-326.

Kocher, T. D., J. A. Conroy, K. A. McKaye, J. R. Stauffer, and S. F. Lockwood. 1995. Evolution of NADH dehydrogenase subunit 2 in East African cichlid fish. Molecular Phylogenetics and Evolution 4:420-432.

Koster, W. J. 1969. Fishes of the Rita Blanca Lake deposits. Pp. 135-139 *in* R. Y. Anderson and D. W. Kirkland, eds. Paleoecology of an Early Pleistocene lake on the High Plains of Texas. Geological Society of America, Boulder, CO.

Kuehne, R. A., and J. W. Small. 1971. *Etheostoma barbouri*, a new darter (Percidae, Etheostomatini) from the Green River with notes on the subgenus *Catonotus*. Copeia 1971:18-26.

Larget, B., and D. L. Simon. 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Molecular Biology and Evolution 16:750-759.

Lindsay, E. H., N. M. Johnson, and N. D. Opdyke. 1975. Preliminary correlation of North American land mammal ages and geomagnetic chronology. Papers on Paleontology from the Museum of Palenotology, University of Michigan 12:111-119.

Liu, H. P., R. Hershler, and K. Clift. 2003. Mitochondrial DNA sequences reveal extensive cryptic diversity within a western American springsnail. Molecular Ecology 12:2771-2782.

Lundberg, J. G., M Kottelat, G. R. Smith, M. L. J. Stiassny, and A. C. Gill. 2000. So many fishes, so little time: an overview of recent ichthyological discovery in continental waters. Annals of the Missouri Botanical Garden 87:26-62.

Maddison, D. R., and W. P. Maddison. 2002. MacClade: analysis of phylogeny and character evolution. Sinauer, Sunderland, MA.

Matthew, W. D. 1924. Third contribution to the Snake Creek fauna. Bulletin of the American Museum of Natural History 50:59-210.

Mayden, R. L. 1985. Biogeography of Ouachita Highland fishes. Southwestern Naturalist 30:195-211.

Mayden, R. L. 1987a. Historical ecology and North American highland fishes: a research program in community ecology. Pp 210-222 *in* W. J. Matthews and D. C. Heins, eds. Community and evolutionary ecology of North American stream fishes. University of Oklahoma Press, Norman, OK.

Mayden, R. L. 1987b. Pleistocene glaciation and historical biogeography of North American central-highland fishes. Pp. 141-152 *in* C. Johnson, ed. Quaternary environments of Kansas. Kansas Geological Survey, Lawrence, KS.

Mayden, R. L. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. Systematic Zoology 37:329-355.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambrige, MA.

Mendelson, T. C., and J. N. Simons. 2006. AFLPs resolve cytonuclear discordance and increase resolution among barcheek darters. Molecular Phylogenetics and Evolution 41:445-453.

Moritz, C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. Molecular Ecology 3:401-411.

Near, T. J., and M. F. Benard. 2004. Rapid allopatric speciation in logperch darters (Percidae: *Percina*). Evolution 58:2798-2808

Near, T. J., and M. J. Sanderson. 2004. Assessing the quality of molecular divergence time estimates by fossil calibrations and fossil-based model selection. Philosophical Transactions of the Royal Society of London, Series B 359:1477-1483.

Near, T. J., and B. P. Keck. 2005. Dispersal, vicariance, and timing of diversification in *Nothonotus* darters. Molecular Ecology 14:3485-3496.

Near, T. J., J. C. Porterfield, and L. M. Page. 2000. Evolution of cytochrome *b* and the molecular systematics of *Ammocrypta* (Percidae: Etheostomatinae). Copeia 2000:701-711.

Near, T. J., L. M. Page, and R. L. Mayden. 2001. Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. Molecular Ecology 10:2235-2240.

Near, T. J., T. W. Kassler, J. B. Koppelman, C. B. Dillman, and D. P. Philipp. 2003. Speciation in North American black basses, *Micropterus* (Actinopterygii: Centrarchidae). Evolution 57:1610-1621.

Near, T. J., D. I. Bolnick, and P. C. Wainwright. 2004. Investigating phylogenetic relationships of sunfishes and black basses (Actinopterygii: Centrarchidae) using DNA sequences from mitochondrial and nuclear genes. Molecular Phylogenetics and Evolution 32:344-357.

Near, T. J., D. I. Bolnick, and P. C. Wainwright. 2005. Fossil calibrations and molecular divergence time estimates in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1768-1782.

Nelson, J. S., E. J. Crossman, H. Espinoza-Perez, L. T. Findley, C. R. Gilbert, R. N. Lea, and J. D. Williams. 2004. Common and scientific names of fishes from the United States, Canada, and Mexico, 6th ed. American Fisheries Society Special Publication 29, Bethesda, MD.

Page, L. M. 1974. The life history of the spottail darter, *Etheostoma squamiceps*, in Big Creek, Illinois and Ferguson Creek, Kentucky. Illinois Natural History Survey Biological Notes 89:1-20.

Page, L. M. 1975. Relations among the darters of the subgenus *Catonotus* of *Etheostoma*. Copeia 1975:782-784.

Page, L. M. 1983. Handbook of darters. T. F.H. Publications, Inc., Neptune City, New Jersey.

Page, L. M. 1985. Evolution of reproductive behaviors in percid fishes. Illinois Natural History Survey Bulletin 33:275-295.

Page, L. M. 2000. Etheostomatinae. Pp. 225-253 *in* J. F. Craig ed. Percid fishes: systematics, ecology, and exploitation. Blackwell Science, Oxford.

Page, L. M., and M. E. Braasch. 1976. Systematic studies of darters of the subgenus *Catonotus* (Percidae), with the description of a new species from the lower Cumberland and Tennessee River systems. Occasional Papers of the Museum of Natural History University of Kansas 60:1-18.

Page, L. M., and H. L. Bart Jr. 1989. Egg mimics in darters (Pisces: Percidae). Copeia 1989:514-518.

Page, L. M., and D. W. Schemske. 1978. The effect of interspecific competition on the distribution and body size of darters of the subgenus *Catonotus* (Percidae: *Etheostoma*). Copeia 1978:406-412.

Page, L. M., P. A. Ceas, D. L. Swofford, and D. G. Buth. 1992. Evolutionary relationships within the *Etheostoma squamiceps* complex (Percidae: subgenus *Catonotus*) with descriptions of five new species. Copeia 1992:615-646.

Page, L. M., M. Hardman, and T. J. Near. 2003. Phylogenetic relationships of barcheek darters (Percidae: *Etheostoma*, subgenus *Catonotus*) with descriptions of two new species. Copeia 2003:512-530.

Passey, B. H., T. E. Cerling, M. E. Perkins, M. R. Voorhies., J. M. Harris, and S. T. Tucker. 2002. Environmental change in the Great Plains: an isotopic record from fossil horses. Journal of Geology 110:123-140.

Pera, T. P., and J. W. Armbruster. A new species of *Notropis* (Cypriniformes: Cyprinidae) from the southeastern United States. Copeia 2006:423-430.

Pflieger, W. L. 1971. A distributional study of Missouri fishes. University of Kansas Publications Museum of Natural History 20:225-570.

Porter, B. A., A. C. Fiumera, and J. C. Avise. 2002. Egg mimicry and allopaternal care: two mate-attracting tactics by which nesting striped darter (*Ethestoma virgatum*) males enhance reproductive success. Behavioral Ecology and Sociobiology 51:350-359.

Porterfield, J. C., L. M. Page, and T. J. Near. 1999. Phylogenetic relationships among fantail darters (Percidae: *Etheostoma*: *Catonotus*): total evidence analysis of morphological and molecular data. Copeia 1999:551-564.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817-818.

Repenning, C. A. 1987. Biochronology of the microtine rodents of the United States. PP. 236-268 *in* M. O. Woodburne, ed. Cenozoic mammals of North America. University of California Press, Berkeley, CA.

Roe, K. J., P. M. Harris, and R. L. Mayden. 2002. Phylogenetic relationships of the genera of North American sunfishes and basses (Percoidei: Centrarchidae) as evidenced by the mitochondrial cytochrome *b* gene. Copeia 2002:897-905.

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

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Molecular Biology and Evolution 19:101-109.

Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19:301-302.

Smith, C. L. 1962. Some Pliocene fishes from Kansas, Oklahoma, and Nebraska. Copeia 1962:505-520.

Smith, G. R. 1981. Late Cenozoic freshwater fishes of North America. Annual Review of Ecology and Systematics. 12:163-193.

Smith, G. R., and J. G. Lundberg. 1972. The Sand Draw fish fauna. Pp. 40-54 *in* M. F. Skinner and C. W. Hibbard, eds. Pleistocene preglacial and glacial rocks and faunas of North Central Nebraska. American Museum of Natural History, New York.

Smith, G. R., and R. R. Miller. 1985. Taxonomy of fishes from Miocene Clarkia Lake beds, Idaho. Pp. 75-83 *in* C. J. Smiley, ed. Late Cenozoic history of the Pacific Northwest

Strange, R. M., and B. M. Burr. 1997. Intraspecific phylogeography of North American highland fishes: a test of the Pleistocene vicariance hypothesis. Evolution 51:885-897.

Switzer, J. F., and R. M. Wood. 2002. Molecular systematics and historical biogeography of the Missouri saddled darter *Etheostoma tetrazonum* (Actinopterygii: Percidae). Copeia 2002:450-455.

Swofford, D. L. 2000. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, MA.

Tedford, R. H., M. F. Skinner, R. W. Fields, J. M. Rensberger, D. P. Whistler, T. Galusha, B. E. Taylor, J. R. Macdonald, and S. D. Webb. 1987. Faunal succession and the biochronology of the Arikareean through Hemphillian interval (Late Oligocene through earliest Pliocene epochs) in North America. Pp. 153-210 *in* M. O. Woodburne, ed. Cenozoic mammals of North America. University of California Press, Berkeley, CA.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 22:4673-4680.

Wilcox, T. P., L. Hugg, J. A. Zeh, and D. W. Zeh. 1997. Mitochondrial DNA sequencing reveals extreme genetic differentiation in a cryptic species complex of neotropical pseudoscorpions. Molecular Phylogenetics and Evolution 7:208-216.

Wiley, E. O., and R. L. Mayden. 1985. Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. Annals of the Missouri Botanical Gardens 72:596-635.

Wilson, R. L. 1968. Systematics and faunal analysis of Lower Pliocene vertebrate assemblage from Trego County, Kansas. Contributions from the Museum of Paleontology, University of Michigan 22:75-126.

Wood, R. M., and M. E. Raley. 2000. Cytochrome *b* sequence variation in the crystal darter *Crystallaria asprella* (Actinopterygii: Percidae). Copeia 2000:20-26.

Zhang, D. and G. M. Hewitt. 2003. Nuclear DNA analysis in genetic studies of populations: practice, problems and prospects. Molecular Ecology 12:563-584.

# APPENDIX

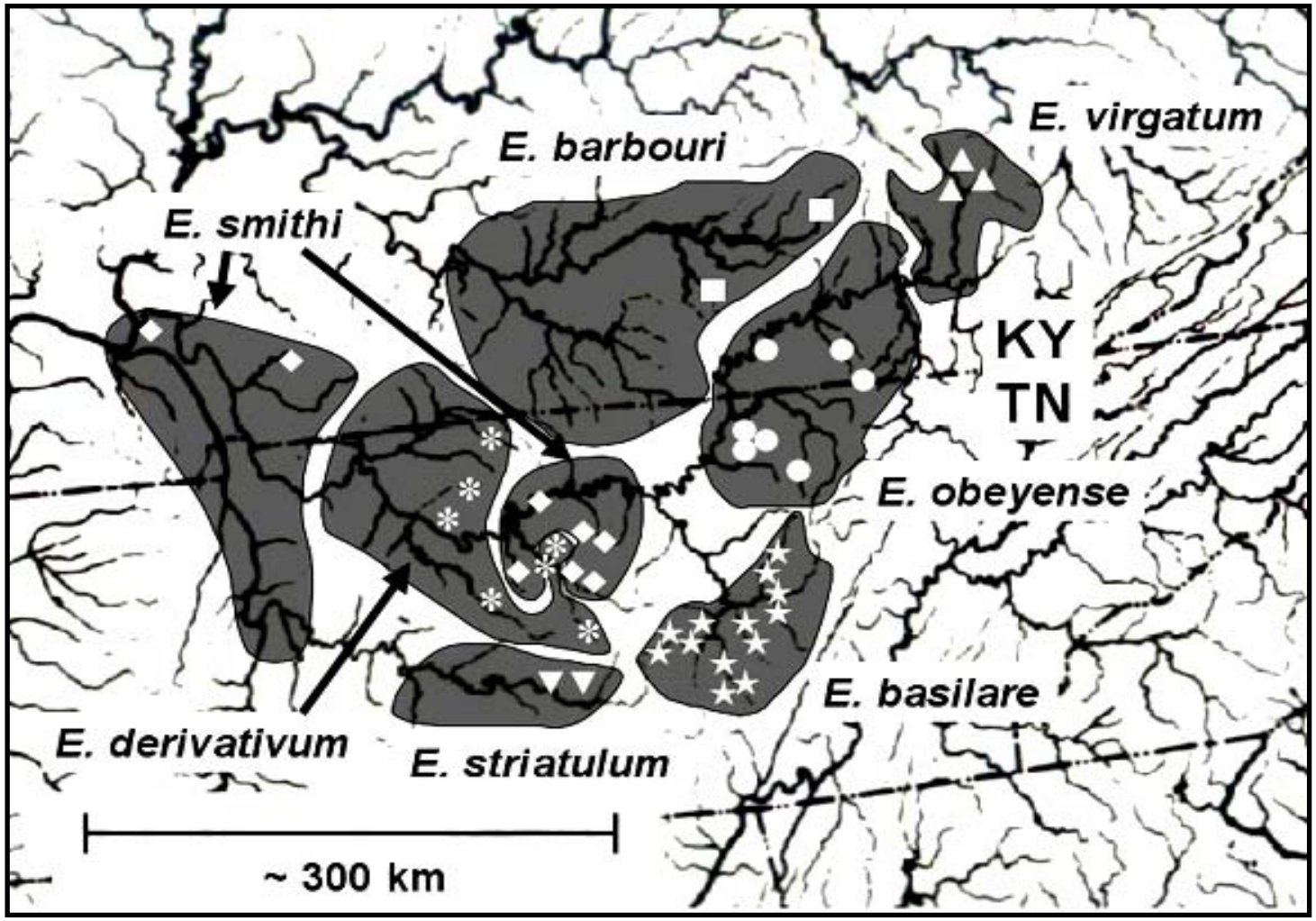


Figure 1. Range map of 7 described species of barcheek darters with sampling localities marked.

Figure 2. *E. obeyense* showing characteristic bar on the operculum.

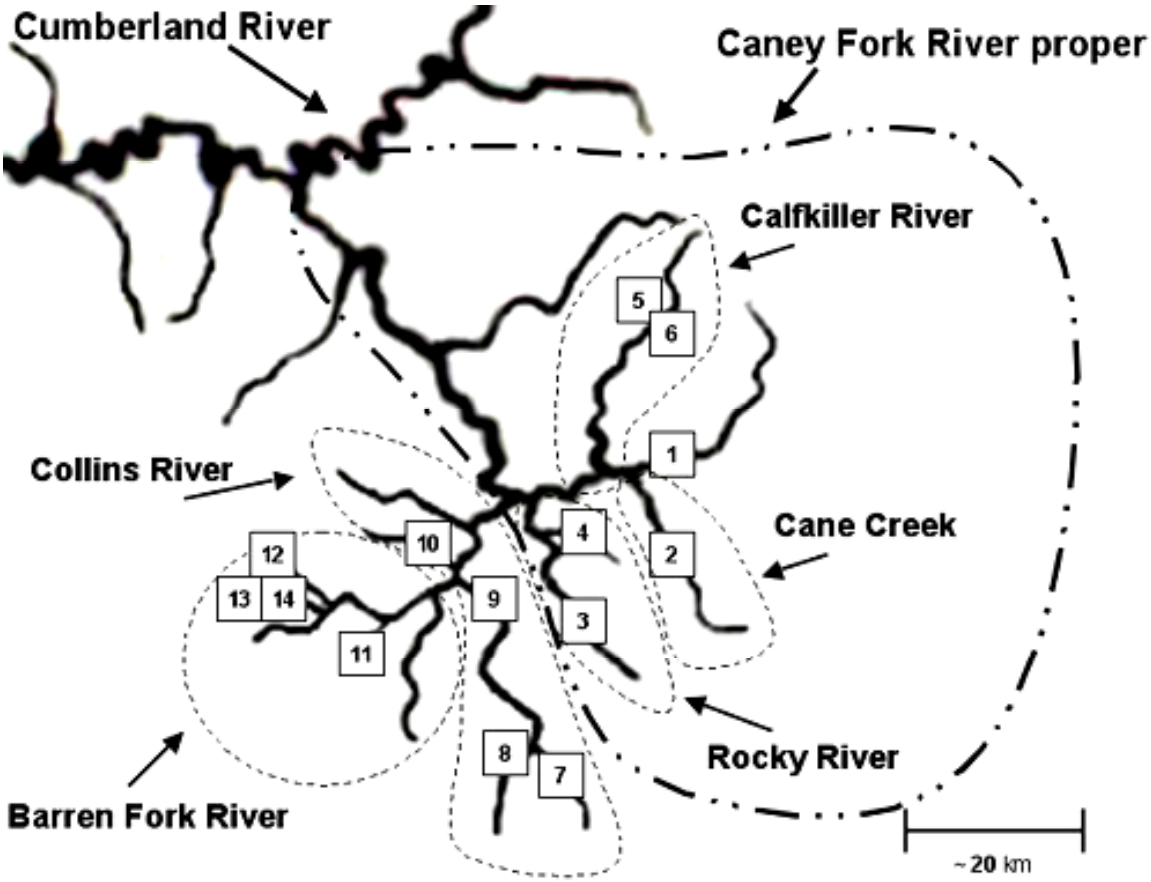


Figure 3. Range of *Etheostoma basilare* with approximate sampling localities marked: 1) upper Caney Fork River 2) Cane Creek at Millstone Branch confluence 3) Rocky River at Boyd's Spring 4) Laurel Creek 5,6) Calfkiller River at Mill Creek confluence 7) Collins River 8) Scott Creek 9) Collins River 10) Charles Creek 11) Garner Branch 12) unnamed tributary to Duke Creek 13) Duke Creek 14) McMahon Creek. See Table 1 for latitude and longitude of localities. Dotted lines correspond to monophyletic clades found in parsimony and Bayesian analysis (Figures 6 and 9, respectively).

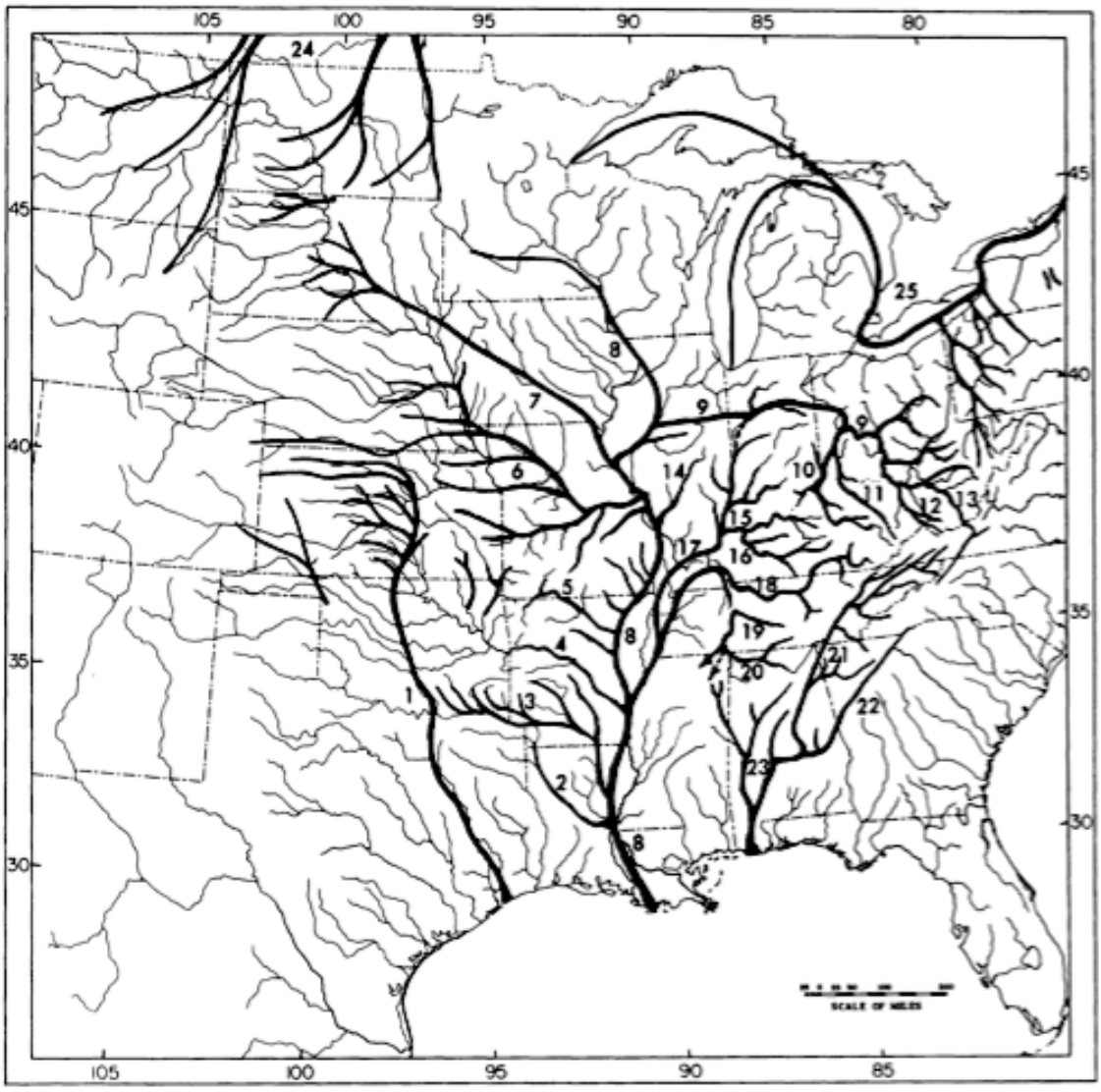


Figure 4. Hypothesized pre-glacial configuration of rivers in eastern North America. 16 is the Green River, 18 is the Old Cumberland River, and 19 is the Old Duck River. (from Mayden 1988)

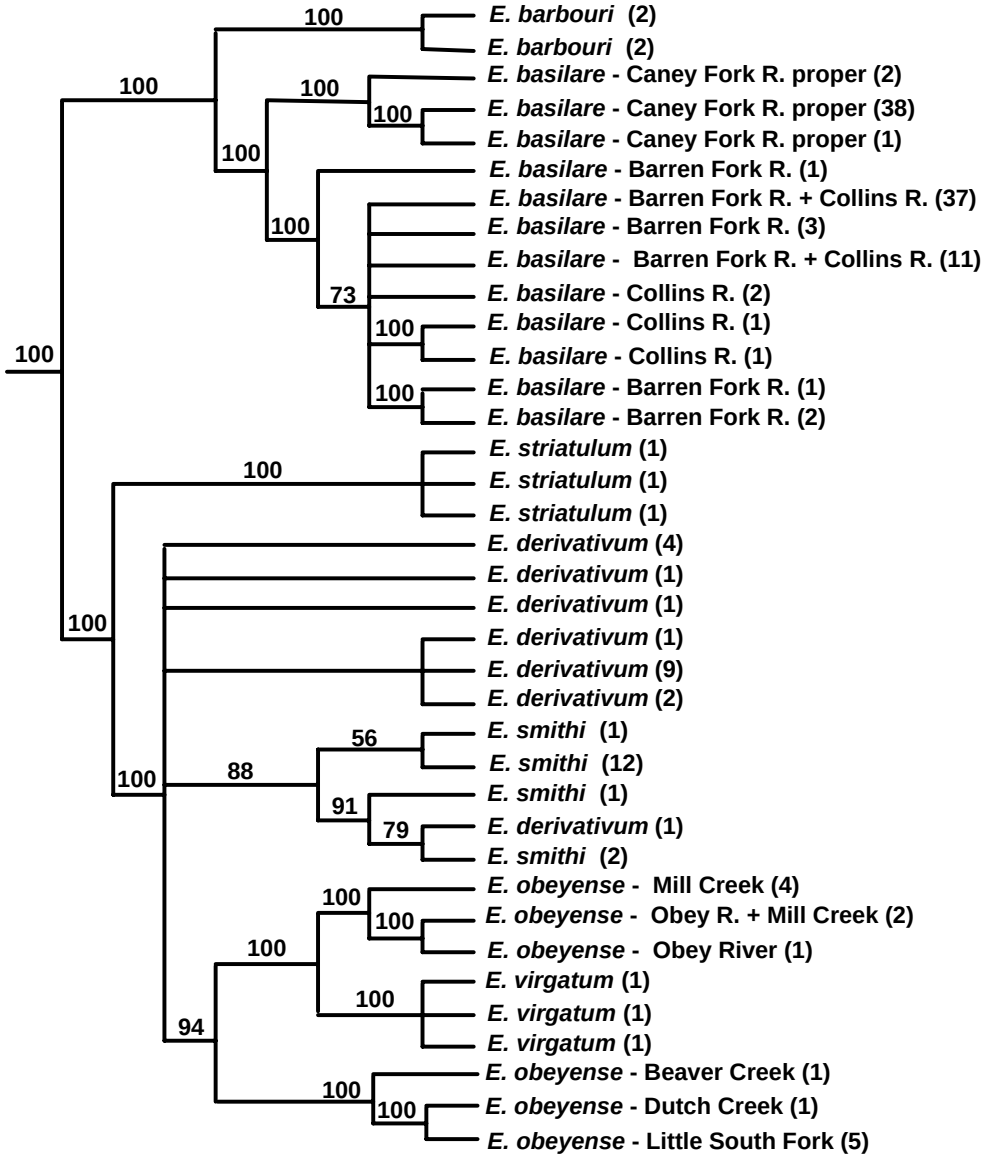


Figure 5. 50% majority-rule consensus tree of 67 most parsimonious trees for 37 unique S7 ribosomal intron alleles sampled from the seven barcheek species. Numbers above branches indicate clade occurrence frequency in the 67 most parsimonious trees. Numbers in parenthesis after taxa names represent the number of individuals sharing that allele. Bootstrap values were all below 75 and are not reported. Nodes receiving a bootstrap score of less than 50 have been collapsed.

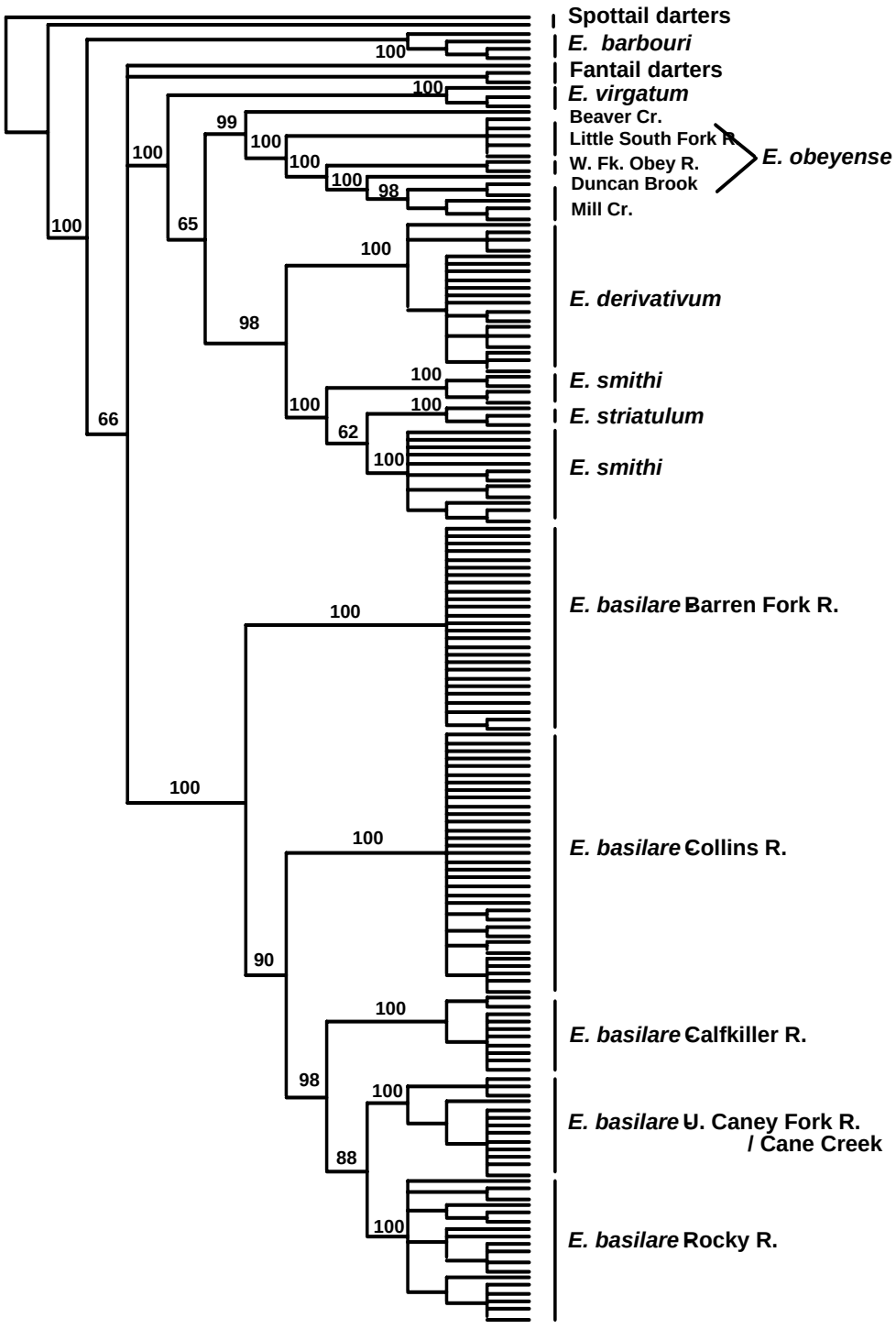


Figure 6. Strict consensus tree resulting from maximum parsimony analysis of 2956 base pairs of mtDNA + nDNA from 159 barcheek darter individuals. The outgroup for this analysis was 5 individuals representing the other two species groups in Catonotus. Numbers above the trees represent bootstrap support values. Nodes receiving a bootstrap score of less than 50 have been collapsed.

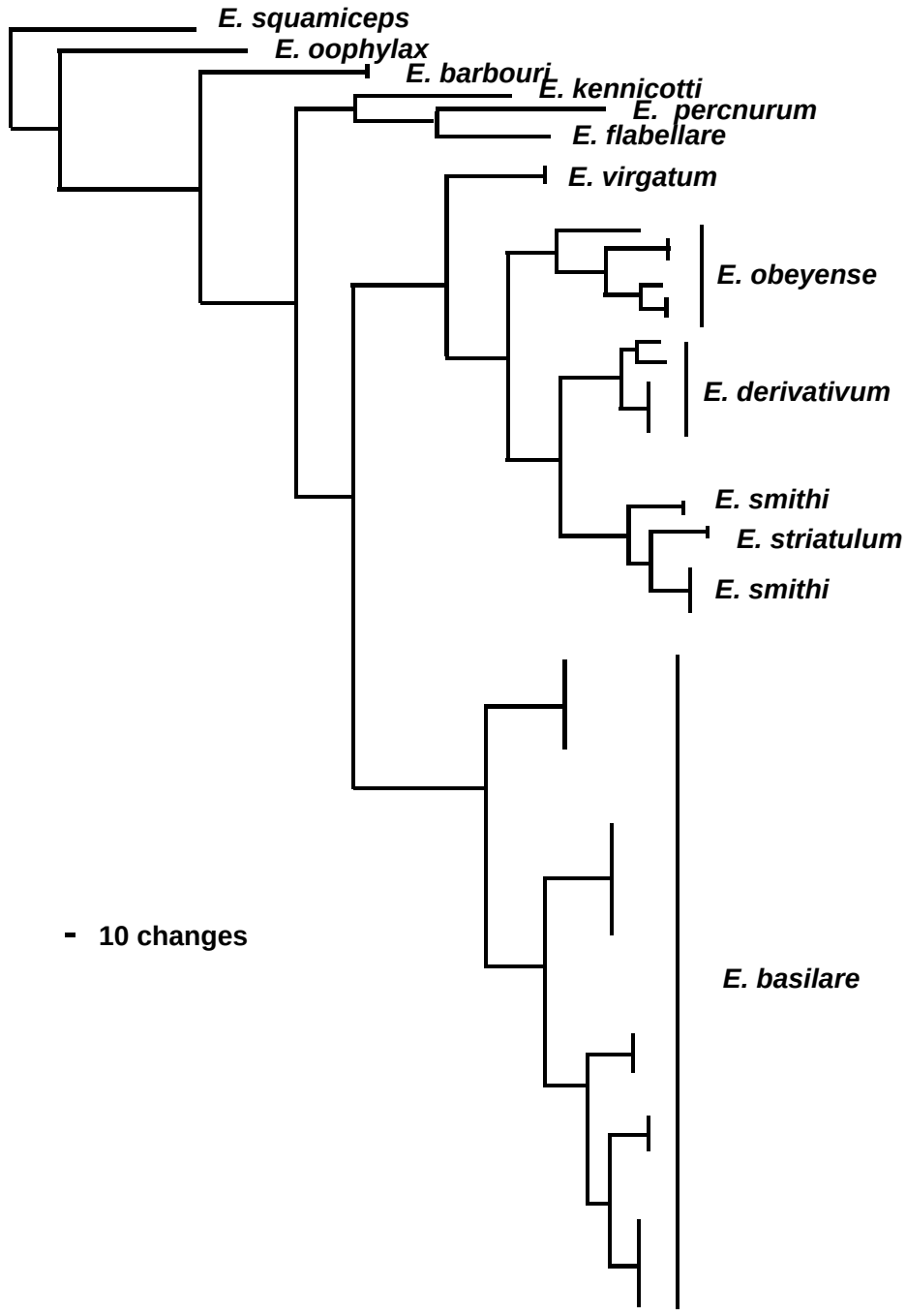


Figure 7. Phylogram representing 1 of 23 most parsimonious trees obtained from analysis of mtDNA + nDNA concatenated data set. Terminal branches less than 10 steps long have been removed.

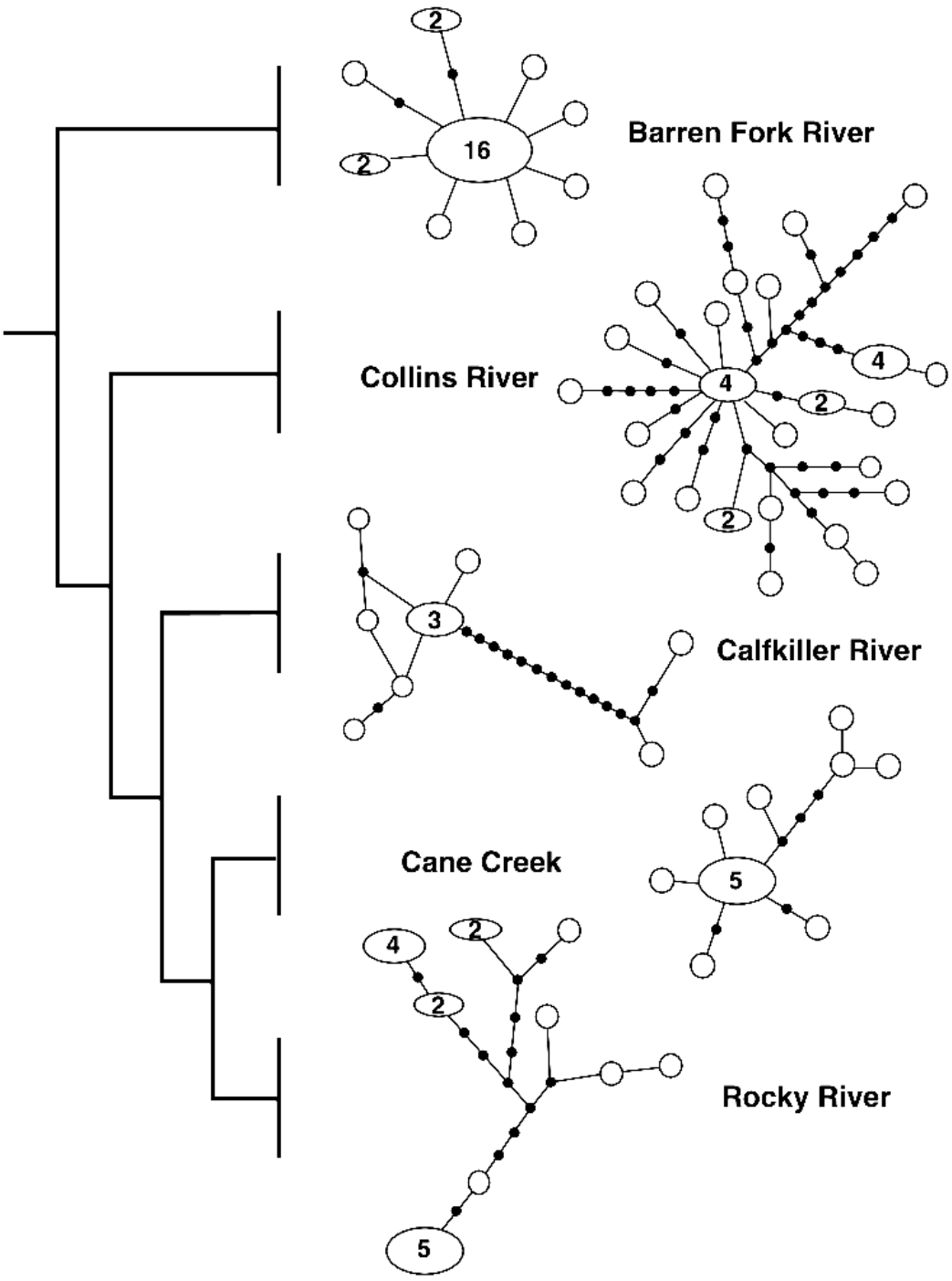


Figure 8. Haplotype networks based on 2397 base pairs of mitochondrial DNA sequence data from 99 individuals of *E. basilare*. Open circles represent unique haplotypes. Numbers in open ovals represent the number of individuals sharing that particular haplotype. Black dots represent inferred missing haplotypes.

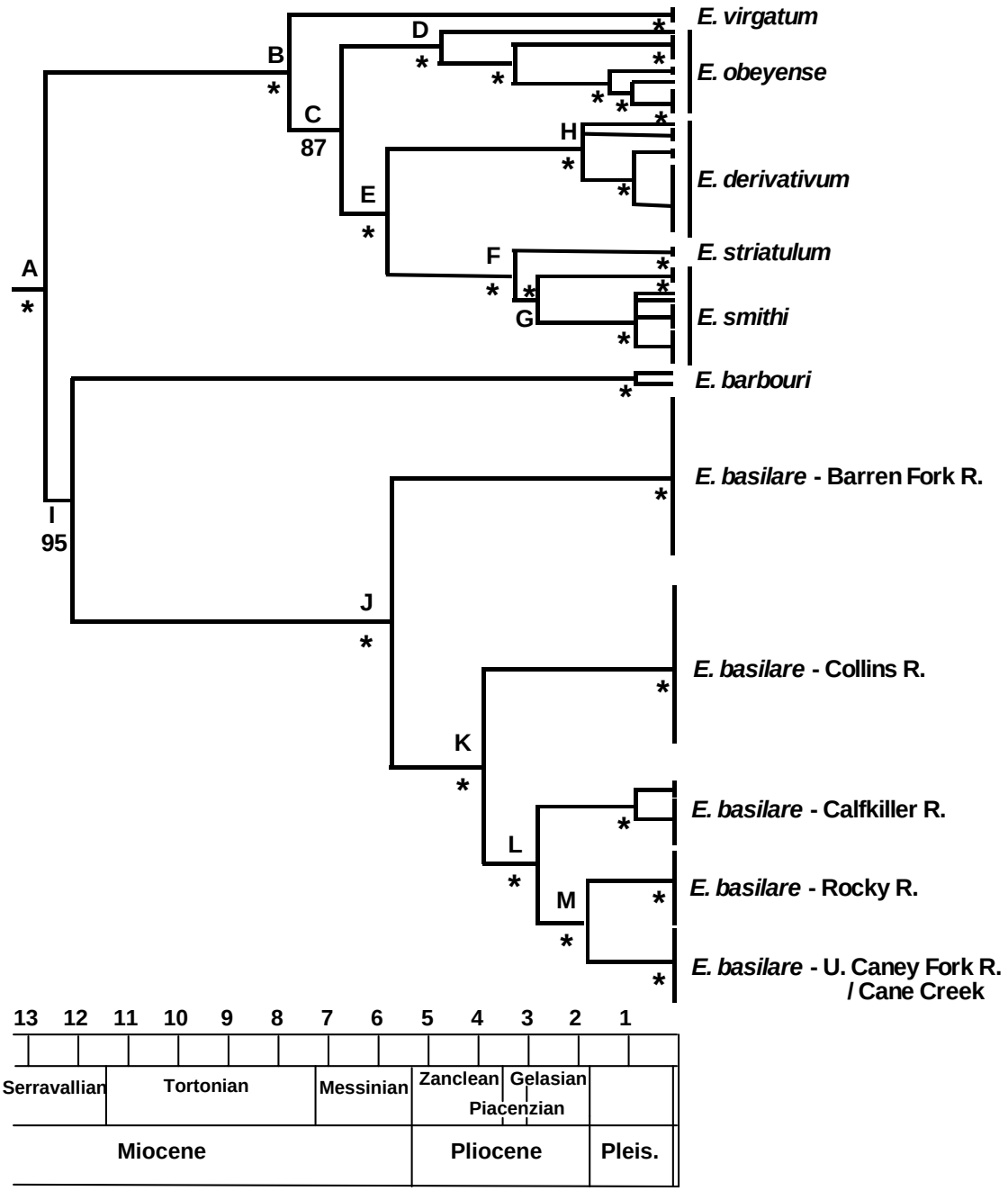


Figure 9. 50 % majority rule consensus chronogram of all trees after the burn-in for the pMM Bayesian analysis. Stars indicate a posterior probability of 100, otherwise posteriors are listed under the letter for each node. The letters correspond to estimated divergence times in Table 5. Geologic time scale is from Gradstein et al. (2004)

Table 1. Specimen collection localities and museum voucher identification numbers. Number in parenthesis after locality indicates the number of individuals sampled from that locality. An asterix after the parentheses indicates previously published data. Genbank accession numbers for these previously published sequences are given in Table 2.

| Species | Locality (number of individuals) | Latitidue, Longitude | Voucher ID number(s) |
|---|---|---|---|
| *Etheostoma barbouri* | Price Creek, Casey Co., KY (2)* | 37°10' 49", 84° 56' 33" | INHS 27864 |
| | East Fork of Little Barren River, Metcalf Co., KY (2) | 37° 0' 41", 85° 32' 44" | no voucher |
| *E. basilare* | tributary to Duke Creek, Cannon Co., TN (2)* | 35° 40' 23", 86° 5' 4" | INHS 27838 |
| | Duke Creek, Cannon Co., TN (9) | 35° 39' 59", 86° 3' 48" | UT 91.6944 |
| | McMahon Creek, Cannon Co., TN (9) | 35° 42' 59", 86° 3' 28" | UT 91.6946 |
| | Garner Branch, Warren Co., TN (6) | 35° 38' 37", 85° 53' 58" | UT 91.7106 |
| | Scott Creek, Warren Co., TN (7) | 35° 34' 15", 85° 42' 40" | UT 91.6704 |
| | Collins River, Warren Co., TN (10) | 35° 40' 30", 85° 42' 35" | UT 91.6940 |
| | Charles Creek, Warren Co., TN (5) | 35° 43' 26", 85° 47' 5" | UT 91.6589 |
| | Rocky River, Van Buren Co., TN (9) | 35° 41' 2", 85° 34' 44" | UT 91.6582 |
| | Laurel Creek, Van Buren Co., TN (9) | 35° 45' 2", 85° 33' 54" | UT 91.6939 |
| | Cane Creek, Van Buren Co., TN (10) | 35° 46' 57", 85° 24' 17" | UT 91.6624 |
| | Caney Fork River, White Co., TN (3) | 35° 49' 5", 85° 21' 29" | UT 91.6592 |
| | Mill Creek, Putnam Co., TN (10) | 36° 5' 16", 85° 53' 58" | UT 91.7022 |
| | Collins River, Grundy Co., TN (10) | 35° 31' 4", 85° 40' 38" | UT 91.6948 |

Table 1. Continued.

| Species | Locality (number of individuals) | Latitidue, Longitude | Voucher ID number(s) |
|---|---|---|---|
| *E. derivativum* | North Fork of Suggs Creek, Wilson Co., TN (3) | 36° 7' 46", 86° 28' 59" | no voucher |
| | Arrington Creek, Williamson Co., TN (2)* | 35° 51' 55", 86° 53' 47" | no voucher |
| | Sycamore Creek, Robertson/Davidson Co., TN (2) | 36° 23' 25", 86° 53' 56" | no voucher |
| | Stones River, Rutherford Co., TN (3) | 35° 48' 20", 86° 25' 29" | no voucher |
| | Hurricane Creek, Rutherford Co., TN (2) | 35° 42' 41", 86°16' 20" | no voucher |
| | Carr Creek, Robertson Co., TN (3) | 36° 28' 19", 86° 54' 53" | no voucher |
| | South Fork of Harpeth River, Williamson Co., TN (3) | 35° 59' 31", 87° 2' 58" | UT 91.7001 |
| *E. obeyense* | Dutch Creek, Cumberland Co., KY (1)* | 36° 49' 6", 85° 26' 54" | INHS 48194 |
| | Duncan Brook, Wayne Co., KY (1) | 36° 45' 41", 84° 52' 31" | no voucher |
| | West Fork of Obey River, Overton Co., TN (2) | 36° 19' 45", 85° 11' 23" | no voucher |
| | Barn Branch of Mill Creek, Overton Co., TN (2) | 36° 27' 5", 85° 22' 33" | UT 91.6690 |
| | Bryan's Fork of Mill Creek, Overton Co., TN (2) | 36° 27' 40", 85° 25' 20" | UT 91.6692 |
| | Morgan Creek, Overton Co., TN (2) | 36° 27' 12", 85° 23' 59" | UT 91.6710 |
| | Little South Fork, Wayne Co., KY (5) | 36° 39' 35", 84° 48' 59" | UT 91.7523 |
| *E. smithi* | Spencer Creek, Wilson Co., TN (3)* | 36° 14' 32", 86° 25' 57" | INHS 52622, UT 91.7084 |
| | Ferguson Creek, Livingston Co., KY (1)* | 37° 8' 28", 88° 21' 37" | INHS 28316 |

Table 1. Continued.

| Species | Locality (number of individuals) | Latitidue, Longitude | Voucher ID number(s) |
|---|---|---|---|
| | Spring Creek, Wilson Co., TN (2) | 36° 5' 18", 86° 13' 51" | no voucher |
| | West Fork of Spring Creek, Wilson Co., TN (2) | 36° 6' 33", 86° 15' 28" | no voucher |
| | East Fork of Stones River, Rutherford Co., TN (2) | 35° 53' 10", 86° 16' 49" | no voucher |
| | Mill Creek, Davidson/Williamson Co., TN (3) | 35° 59' 40', 86° 41' 30" | UT 91.7100 |
| | Muddy Fork of Little River, Chistian Co., KY (1) | 36° 59' 5", 87° 38' 27" | UT 91.7290 |
| *E. striatulum* | Hurricane Creek, Bedford Co., TN (2)* | 35° 32' 25", 86° 27' 7" | NCSM 29833, INHS 48193 |
| | Wartrace Creek, Bedford Co., TN (1) | 35° 32' 35", 86° 20' 30" | no voucher |
| *E. virgatum* | tributary to Roundstone Creek, Rockcastle Co., KY (1)* | 37° 25' 38", 84° 18' 23" | INHS 27832 |
| | Clear Creek, Rockcastle Co., KY (1)* | 37° 28' 16", 84° 15' 19" | INHS 37939 |
| | Middle Fork of Rockcastle River, Jackson, Co. KY (1) | 37° 20' 35", 84° 4' 43" | no voucher |

Table 2. Locality and GenBank accession numbers for all previously published sequences used in this study

| Species | Locality | Genbank Accession Numbers | | |
|---|---|---|---|---|
| | | cyt *b* | ND2 | S7 |
| Barcheeks | | | | |
| *Etheostoma barbouri* | Price Creek, Casey Co., KY (2) | AF412528-29 | AF412542-43 | AF412559-60 |
| *E. basilare* | tributary to Duke Creek, Cannon Co., TN (2) | AF412534<br>AF123043 | AF412548<br>AF412551 | AF412565<br>AF412668 |
| *E. derivativum* | Arrington Creek, Williamson Co., TN (2) | AF412532-33 | AF412549-50 | AF412566-67 |
| *E. obeyense* | Dutch Creek, Cumberland Co., KY | AF123035 | AF412544 | AF412561 |
| *E. smithi* | Ferguson Creek, Livingston Co., KY<br>Spencer Creek, Wilson Co., TN | AF412531<br>AF412530 | AF412546<br>AF412545 | AF412563<br>AF412562 |
| *E. striatulum* | Hurricane Creek, Bedford Co., TN | AF123042 | AF412547 | AF412564 |
| *E. virgatum* | tributary to Roundstone Creek, Rockcastle Co., KY<br>Clear Creek, Rockcastle Co., KY | AF412535<br>AF412536 | AF415552<br>AF412553 | AF412569<br>AF412570 |
| *Catonotus* outgroup | | | | |
| *E. flabellare* | Knights Branch, Vermillion Co., IL | AF412526 | AF412540 | AF412557 |
| *E. kennicotti* | Poor Fork, Letcher Co., KY | AF412527 | AF412541 | AF412558 |
| *E. oophylax* | McCullough Fork, Calloway Co., KY | AF412524 | AF412538 | AF412555 |
| *E. percnurum* | Copper Creek, Scott Co., VA | AF412525 | AF412539 | AF412556 |
| *E. squamiceps* | Big Creek, Hardin Co., IL | AF412523 | AF412537 | AF415554 |

Table 2. Continued.

| Species | Locality | Genbank Accession Numbers | | |
|---|---|---|---|---|
| | | cyt *b* | ND2 | S7 |
| Centrarchidae | | | | |
| *Acantharcus pomotis* | Lake Nummy, Cape May Co., NJ | n/a | AY517726 | AY517757 |
| *Ambloplites ariommus* | Conasauga River, Bradley Co., TN | n/a | AY517727 | AY517758 |
| *Ambloplites cavifrons* | Tar River, Franklin Co., NC | n/a | AY517728 | AY517759 |
| *Ambloplites constellatus* | North Fork White River, Douglas Co., MO | n/a | AY517729 | AY517760 |
| *Ambloplites rupestris* | Lake Andrusia, Beltrami Co., MN | n/a | AY225723 | AY517761 |
| *Archoplites interruptus* | Hume Lake, Fresno Co., CA | n/a | AY225725 | AY517762 |
| *Centrarchus macropterus* | Mud Creek, Hardin Co., TN | n/a | AY225726 | AY517763 |
| *Enneacanthus chaetodon* | Lake Mummy, Cape May Co., NJ | n/a | AY517730 | AY517764 |
| *Enneacanthus gloriosus* | Wacissa River, Jefferson Co., FL | n/a | AY517731 | AY517765 |
| *Enneacanthus obesus* | West Branch Sopchoppy River, Wakulla Co., FL | n/a | AY225724 | AY517766 |
| *Lepomis auritis* | Conasauga River, Bradley Co., TN | n/a | AY517732 | AY517767 |
| *Lepomis cyanellus* | Saline Branch, Champaign Co., IL | n/a | AY517733 | AY517768 |
| | Embarras River, Champaign Co., IL | n/a | AY517734 | AY517769 |
| *Lepomis gibossus* | Lake Andrusia, Beltrami Co., MN | | AY517735 | AY517770 |
| *Lepomis gulosus* | Pine Hills Swamp, Union Co., IL | n/a | AY517736 | AY517771 |

Table 2. Continued.

| Species | Locality | Genbank Accession Numbers | | |
|---|---|---|---|---|
| | | cyt *b* | ND2 | S7 |
| *Lepomis gulosus* | Horsehoe Lake, Alexander Co., IL | | AY517737 | AY517772 |
| *Lepomis humilis* | Mississippi River, Clinton Co., IA | n/a | AY517738 | AY517773 |
| | Horeshoe Lake, Alexander Co., IL | n/a | AY517739 | AY517774 |
| *Lepomis marginatus* | Panther Creek, Henry Co., TN | n/a | AY517741 | AY517777 |
| *Lepomis megalotis* | Saline Branch, Champaign Co., IL | n/a | AY517742 | AY517778 |
| | Horseshoe Lake, Alexander Co., IL | n/a | AY517743 | AY517779 |
| *Lepomis microlophus* | Wacissa River, Jefferson Co., FL | n/a | AY517744 | AY517780 |
| *Lepomis miniatus* | Conasauga River, Bradley Co., TN | n/a | AY225728 | AY517781 |
| | San Marcos River, Hays Co., TX | n/a | AY517745 | AY517782 |
| *Lepomis punctatus* | Wacissa River, Jefferson Co., FL | n/a | AY517746 | AY517783 |
| *Lepomis symmetricus* | Pine Hills Swamp, Union Co., IL | n/a | AY517747 | AY517784 |
| *Micropterus cataractae* | Flint River, Crisp Co., GA | n/a | AY225776 | AY517785 |
| *Micropterus coosae* | Conasauga River, Polk Co., TN | n/a | AY225728 | AY517786 |
| *Micropterus dolomieu* | Fox River, Kenosha Co., WI | n/a | AY225747 | AY517787 |
| | Sugar Creek, MacDonald Co., MO | n/a | AY225751 | AY517788 |
| *Micropterus floridanus* | Lake Eustis, Lake Co., FL (2) | n/a | AY225729-30 | AY517789-90 |
| *Micropterus notius* | Wacissa River, Jefferson Co., FL | n/a | AY225764 | AY517791 |
| | Santa Fe River, Alachua Co., FL | n/a | AY225766 | AY517792 |

Table 2. Continued.

| Species | Locality | Genbank Accession Numbers | | |
|---|---|---|---|---|
| | | cyt *b* | ND2 | S7 |
| *Micropterus punctulatus* | Chase Lake, Chase Co., KS | n/a | AY225755 | AY517793 |
| | Lake Whitney, Hill Co., TX | n/a | AY225761 | AY517794 |
| *Micropterus salmoides* | Lipset Lake, Burnett Co., WI | n/a | AY225735 | AY517795 |
| *Pomoxis annularis* | North Fork White River, Douglas Co., MO (2) | n/a | AY517748 | AY517798 |
| | | n/a | AY517749 | AY517799 |
| *Pomoxis nigromaculatus* | Mud Creek, Hardin Co., TN | n/a | AY517750 | AY517800 |
| | Horseshoe Lake, Alexander Co., IL (2) | n/a | AY517751 | AY517801 |
| | | n/a | AY517752 | AY517802 |

Table 3. Centrarchid fossils and age data used in divergence time analysis of the barcheek darter clade

| Fossil | Age (mya) | Reference | Source for age | Age constrained or fixed |
|---|---|---|---|---|
| *Micropterus spp.* | 16.0 | Matthew 1924 | Tedford et al. 1987 | minimal age constraint |
| *Archoplites clarki* | 15.5 | Smith and Miller 1985 | Golenberg et al. 1990 | fixed |
| *Pomoxis sp.* | 12.0 | Wilson 1968 | Wilson 1968; Tedford et al. 1987 | fixed |
| *Lepomis kansasensis* | 6.6 | Hibbard 1936 | Passey et al. 2002 | fixed |
| *Lepomis humilis* | 3.4 | Smith and Lundberg 1972 | Repenning 1987 | minimal age constraint |
| *Lepomis megalotis* | 2.4 | Koster 1969 | Lindsay et. al 1975; Repenning 1987 | minimal age constraint |

Table 4. Models chosen by ModelTest for the pMM Bayesian and likelihood analyses

| Data partition | Model | Substitution types | Invariant Sites | Substitution Rates |
|---|---|---|---|---|
| ND2 1st codon | GTR | 6 | yes | gamma distributed |
| ND2 2nd codon | GTR | 6 | yes | gamma distributed |
| ND2 3rd codon | GTR | 6 | yes | gamma distributed |
| S7 | TrN | 2 | no | gamma distributed |
| ND2 + S7 | GTR | 6 | yes | gamma distributed |

Table 5. Divergence time estimates for the barcheek. Letters for nodes correspond to Figure 9.

| Node | Estimated age (mya) |
|---|---|
| A | 12.68 |
| B | 7.49 |
| C | 6.71 |
| D | 4.56 |
| E | 5.35 |
| F | 3.09 |
| G | 2.57 |
| H | 1.40 |
| I | 11.87 |
| J | 5.61 |
| K | 3.70 |
| L | 2.47 |
| M | 1.49 |

## Vita

Phillip R. Hollingsworth Jr. was born on October 14th, 1980 in Nashville, TN. He attended grade school at St. Edward School in Nashville and then later moved to Branson, MO were he attended middle and high school. After graduating from Branson High School in 1999 he received a B. S. in biology from Missouri State University.

Phillip currently resides with his girlfriend of two years, Sarah Farnsley, and their bull terrier Mia in Knoxville, TN. Phillip spends his time pursuing his interests in fly fishing, aquatic entomology, and conservation biology.