12-2008

Phylogeography of the Bigeye Chub *Hybopsis amblops* (Teleostei: Cypriniformes): Early Pleistocene Diversification and Post-Glacial Range Expansion

Peter B. Berendzen
University of Northern Iowa

Tony Gamble
Marquette University, anthony.gamble@marquette.edu

Andrew M. Simons
University of Minnesota

Journal of Fish Biology, Vol. 73, No. 8 (December 2008): 2021-2039. DOI.

Phylogeography Of the Bigeye Chub *Hybopsis Amblops* (Teleostei: Cypriniformes): Early Pleistocene Diversification and Post-Glacial Range Expansion

P. B. Berendzen
Department of Fisheries, Wildlife, and Conservation Biology and Bell Museum of Natural History, University of Minnesota, 1987 Upper Buford Circle, St Paul, MN
Conservation Biology Graduate Program and Bell Museum of Natural History, University of Minnesota, 1987 Upper Buford Circle, St Paul,MN
T. Gamble
Department of Fisheries, Wildlife, and Conservation Biology and Bell Museum of Natural History, University of Minnesota, 1987 Upper Buford Circle, St Paul, MN
Conservation Biology Graduate Program and Bell Museum of Natural History, University of Minnesota, 1987 Upper Buford Circle, St Paul, MN
A. M. Simons

Department of Fisheries, Wildlife, and Conservation Biology and Bell Museum of Natural History, University of Minnesota, 1987 Upper Buford Circle, St Paul, MN
Conservation Biology Graduate Program and Bell Museum of Natural History, University of Minnesota, 1987 Upper Buford Circle, St Paul, MN

Abstract

The bigeye chub, *Hybopsis amblops*, is a member of the Central Highlands ichthyofauna of eastern North America. Phylogenetic analyses of the *H. amblops* species group based on a 1059 bp fragment of the mitochondrial DNA cytochrome *b* gene did not recover a monophyletic group. The inclusion of *Hybopsis hypsinotus* in the species complex is questionable. Within *H. amblops*, five strongly supported clades were identified; two clades containing haplotypes from the Ozark Highlands and three clades containing haplotypes from the Eastern Highlands and previously glaciated regions of the Ohio and Wabash River drainages. Estimates of the timing of divergence indicated that prior to the onset of glaciation, vicariant events separated populations east and west of the Mississippi River. East of the Mississippi River glacial cycles associated with the blocking and rerouting of the Teays River system caused populations to be pushed southward into refugia of the upper Ohio River. Following the most recent Wisconsinan glaciation, populations expanded northward into previously glaciated regions and southward into the Cumberland River drainage. In the Ozarks, west of the Mississippi River, isolation of clades appears to be maintained by the lack of stream capture events between the upper Arkansas and the White River systems and a barrier formed by the Arkansas River.

Introduction

Three major upland areas separated by intervening lowlands comprise the Central Highlands of eastern North America, Ouachita, Ozark and Eastern highlands (**Fig. 1**). These areas are characterized by clear, cool, high-gradient streams and have similarities in their fauna sharing species and species groups. The Mississippi River is the primary centre of diversity and distribution of the eastern North American ichthyofauna (**Burr & Page, 1986; Robison, 1986**). The Central Highlands contain a major portion of this diversity, which exhibits a replicated pattern of disjunct distributions in each of the highland regions (**Mayden, 1988**). There are a number of widespread species that are shared among highland regions and have subsequently dispersed into previously glaciated regions, *e.g. Etheostoma caeruleum* Storer (**Burr & Page, 1986**), *Percina evides* (Jordan & Copeland) (**Near *et al.*, 2001**), *Erimystax dissimilis* (Kirtland) and *Erimystax x-punctatus* (Hubbs & Crowe) (**Simons, 2004**) and *Hypentelium nigricans* (Lesueur) (**Berendzen *et al.*, 2003**). There are also closely related species groups with endemics in each highland region, *e.g. Etheostoma variatum* Kirtland species group (**Wiley & Mayden, 1985**), *Micropterus* species group (**Near *et al.*, 2003**) and *Notropis rubellus* (Agassiz) species group (**Berendzen *et al.*, 2008**).

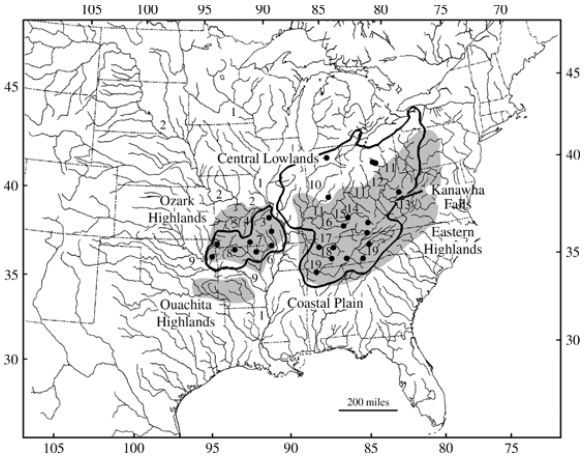

Figure 1 Map of the Eastern U.S.A. with pertinent rivers indicated and highland areas shaded. 1, Mississippi River; 2, Missouri River; 3, Meramec River; 4, Gasconade River; 5, Osage River; 6, St Francis River; 7, Black River Drainage; 8, White River Drainage; 9, Arkansas River; 10, Wabash River Drainage; 11, Ohio River; 12, Muskingum River; 13, Kanawha River; 14, Kentucky River; 15, Salt River; 16, Green River; 17, Cumberland River; 18, Duck River and 19, Tennessee River. Bold outline represents the approximate distribution of *Hybopsis amblops*. Black dots identify sampling localities.

The Central Highlands were influenced by periodic climatic oscillations during the Quaternary. The terms Pleistocene and Quaternary are associated with the Ice Age, the period of time marked by glacial and interglacial cycles (**Gibbard & van Klofschoten, 2004**). The Quaternary is traditionally considered to be the interval of oscillating climatic extremes with cold or glacial stages interspersed with warm or interglacial periods (**Lowe & Walker, 1997**) that was initiated *c.* 2·6 million years ago (MYA) at the beginning of the Gelasian of the late Pliocene and extending to the present (**Gibbard & van Klofschoten, 2004**). The date of Pliocene–Pleistocene boundary is controversial; there is an ongoing debate as to whether the base of the Pleistocene should be placed at the first evidence of climatic cooling *c.* 2·6 MYA (**Gibbard & van Klofschoten, 2004**). However, it has most recently been defined as 1·81 MYA based on Global Stratotype Section and Point (**Lourens *et al.*, 2004**). Consequently, the date of this boundary is no longer associated with climatic change.

Prior to the onset of climatic change and glacial cycles (*c.* 2·6 MYA), the Central Highlands were thought to be a large contiguous highland region (**Thornbury, 1965**). The region was subsequently fragmented by a series of events associated with advancing and retreating glacial fronts causing major changes in river flow and character and drainage patterns (**Fisk, 1944; Thornbury, 1965; Mayden, 1988**). The continental ice sheets in North America advanced in stages and varied in size and maximum extension of the southern boundary (**Burr & Page, 1986**). West of the Mississippi River, the earliest pre-Illinoian glacial advances, constituting at least 10 stages from the early Gelasian of the late Pliocene (2·6 MYA) to the middle Pleistocene (0·34 MYA) (**Lowe & Walker, 1997; Mickelson & Colgan, 2004**), extended well into Northern Missouri terminating approximately along the course of the modern Missouri River (**Burr & Page, 1986**). The younger Illinoian (0·08–0·01 MYA) and the most recent Wisconsinan (0·30–0·13 MYA) glacial cycles (**Lowe & Walker, 1997**) did not extend as far southward west of the Mississippi River and had less of an effect on the Ozark Highlands. Thus, the drainage patterns of the Ozark Highlands were not directly affected by glaciation and have maintained their basic configuration (**Pflieger, 1971**). However, the character of the streams has changed over

time. East of the Mississippi River, the pre-Illinoian, Illinoian and most recent Wisconsinan glacial cycles of the Quaternary had a maximum extension of the southern boundary much farther south than west of the Mississippi River, reaching as far south as the modern Ohio River (**Burr & Page, 1986**). Like the Ozark Highlands, the drainage patterns in the unglaciated regions of the Eastern Highlands south of the Modern Ohio River were not directly affected by glaciation. Today, the Central Highlands include the Ozark and Ouachita Highlands west of the Mississippi River and the Eastern Highlands containing the Appalachian Mountains east of the Mississippi River (**Fig. 1**).

Hybopsis amblops (Rafinesque), the bigeye chub, is a member of the Central Highlands fauna with a widespread, disjunct distribution in the Mississippi River and Laurentian Great Lakes basins (**Fig. 1**). West of the Mississippi River, it is found in portions of the Ozark Highlands including Meramec River of the northern Ozarks, the St Francis and White rivers in the southern Ozarks and tributaries of the middle Arkansas River in the western Ozarks. East of the Mississippi River, it is found in previously glaciated regions of the Ohio River and Great Lakes drainages and in the unglaciated regions of the Eastern Highlands (**Fig. 1**). In the previously glaciated regions of the Ohio River drainage, it is found in portions of the Illinois River drainage, Kaskaskia and Wabash River drainages and tributaries of the upper Ohio River draining southward from the Great Lakes region. In the unglaciated regions of the Eastern Highlands, it is found in tributaries of the upper Ohio River draining northward including the Cumberland and Tennessee rivers.

Hybopsis amblops lives in clear, cool to warm, moderate-gradient streams and rivers. They are typically found in flowing pools and backwaters, slow runs and often along *Justicia* beds (**Jenkins & Burkhead, 1994**). This fish is rarely found in turbid water, and their presence is often considered an indicator of excellent water quality (**Boschung & Mayden, 2004**). Populations have declined in the northern part of the species' range, and it is listed as possibly extirpated in Michigan and critically imperilled in Illinois (NatureServe, **http://www.natureserve.org**). *Hybopsis amblops* is included in the *H. amblops* species group, which contains *H. amblops*, *Hybopsis amnis* (Hubbs & Greene), *Hybopsis hypsinotus* (Cope), *Hybopsis lineapunctata* Clemmer & Suttkus, *Hybopsis rubrifrons* (Jordan), *Hybopsis winchelli* Girard and the undescribed form *Hybopsis* sp., *cf*. *winchelli* (**Shaw *et al.*, 1995; Grose & Wiley, 2002**).

The objective of this study was to examine genetic variation across the range of *H. amblops* and use these patterns to assess the roles of vicariance and population expansion in shaping the present diversity and distribution within the group. Previous hypotheses explained the disjunct distribution of *H. amblops* in the Central Highlands in the context of a dispersalist scenario. **Burr & Page (1986)** suggested that populations were separated into multiple refugia east and west of the Mississippi River by early glacial advance and subsequently expanded into previously glaciated regions following glacial retreat.

Phylogenetic and demographic patterns of *H. amblops* were inferred from a fragment of the mitochondrial gene cytochrome *b* (cyt *b*). The following hypotheses were tested: (1) the *H. amblops* species group is monophyletic; (2) at the onset of glaciation, *H. amblops* was separated into refugia east and west of the Mississippi River and following glacial retreat expanded into previously glaciated regions; (3) the demographic patterns within *H. amblops* are consistent with hypotheses of the timing of geological events during the Quaternary east and west of the Mississippi River and (4) the

demographic patterns within *H. amblops* are consistent with hypotheses of historical drainage patterns in eastern North America. Finally, the results are compared with previously published phylogeographical studies of Central Highlands fishes.

Materials and methods

Sampling

Specimens of *H. amblops* were collected with seines and backpack electroshocker from 24 localities across its range (**Fig. 1 and Table I[link]**). Fishes were frozen in liquid nitrogen and transported to the laboratory where they were stored in an ultracold freezer at −80° C. *Hybopsis amnis*, *H. hypsinotus*, *H. lineapunctata*, *H. rubrifrons* and *H. winchelli*, members of the hypothesized *H. amblops* species group (**Grose & Wiley, 2002**), were included as outgroups. The putative undescribed form of *H. winchelli* from east of the Mobile Basin was not included in the analyses. Additionally, more distantly related outgroup taxa were also included, e.g. *Luxilus coccogenis* (Cope) and *Notropis telescopus* (Cope) (**Table I**).

Table I. Specimens used in this study listed with the population name, number of specimens from each locality in parentheses, catalogue number, GenBank accession number and collection locality

Hybopsis amblops	Salt (5), JFBM 38430, EU917371–EU917375, North Rolling Fork River, Boyle Co., KY, U.S.A.
	Kentucky (2), JFBM uncat., EU917359 and EU917360, South Fork Kentucky River, Owsley Co., KY, U.S.A.
	MuskingumA (1), JFBM uncat., EU917386, Walhonding River, Coshocton Co., OH, U.S.A.
	MuskingumB (5), JFBM 38296, EU917387–EU917391, Walhonding River, Coshocton Co., OH, U.S.A.
	Kanawha (5), JFBM 42544, EU917354–EU917358, Elk River, Braxton Co., WV, U.S.A.
	Green (5), JFBM 38209, EU917342–EU917346, Green River, Green Co., KY, U.S.A.
	CumberlandA (5), JFBM 34907, EU917329–EU917333, East Fork Stones River, Rutherford Co., TN, U.S.A.
	CumberlandB (3), JFBM 40560, EU917334–EU917336, Turnbull Creek, Cheatam Co., TN, U.S.A.
	Duck (5), JFBM 35784, EU917337–EU917341, Duck River, Bedford Co., TN, U.S.A.
	Shoal (5), JFBM 40964, EU917376–EU917380, Shoal Creek, Lauderdale Co., AL, U.S.A.
	Sequatchie (5), JFBM 35548, EU917366–EU917370, Sequatchie River, Marion Co., TN, U.S.A.
	Powell (5), JFBM 35127, EU917392–EU917396, Powell River, Claiborne Co., TN, U.S.A.

	Little (2), JFBM uncat., EU917397 and EU917398, Little River, Blount Co., TN, U.S.A.
	Hiwassee (7), JFBM uncat., EU917347–EU917353, Valley River, Cherokee Co., NC, U.S.A.
	WhiteA (5), JFBM 39341, EU917408–EU917412, Kings River, Carroll Co., AR, U.S.A.
	WhiteB (5), JFBM 39471, EU917413–EU917417, North Fork White River, Ozark Co., MO, U.S.A.
	BlackA (4), JFBM 40888, EU917320–EU917323, Strawberry River, Sharp Co., AR, U.S.A.
	BlackB (5), JFBM 30470, EU917324–EU917328, Current River, Clay Co., AR, U.S.A.
	St Francis (5), JFBM uncat., EU917381–EU917385, St Francis River, Madison Co., MO, U.S.A.
	Meramec (5), JFBM uncat., EU917361–EU917365, Big River, Washington Co., MO, U.S.A.
	ArkansasA (1), JFBM uncat., EU917316, Elk River, McDonald Co., MO, U.S.A.
	ArkansasB (3), JFBM uncat., EU917317–EU917319, Illinois River, Cherokee Co., OK, U.S.A.
	WabashA (4), JFBM 34853, EU917399–EU917402, Tippecanoe River, Fulton Co., IN, U.S.A.
	WabashB (5), JFBM 35391, EU917403–EU917407, East Fork White River, Jackson Co., IN, U.S.A.
Hybopsis amnis	(5), JFBM uncat., EU917418–EU917422, Hatchie River, Hardeman Co., TN, U.S.A.
Hybopsis hypsinotus	A UAIC 12487.01, EU917423, Pauls Creek, Surry Co., NC, U.S.A.
	B UAIC 7930.02, EU917424, Turkey Creek, Chester Co., SC, U.S.A.
Hybopsis lineapunctata	(4), JFBM 35187, EU917425–EU917428, Josie Leg Creek, Tallapoosa Co., AL, U.S.A.
Hybopsis rubrifrons	(5), UAIC 7921.02, EU917429–EU917433, Candler Creek, Hall Co., GA, U.S.A.
Hybopsis winchelli	A (4), JFBM 35843, EU917436–EU917439, East Fork Amite River, Amite Co., MS, U.S.A.
	B (2), JFBM 37608, EU917434 and EU917435, Cahaba River, Bibb Co., AL, U.S.A.
Luxilus coccogenis	(1), ASU uncat., U66603
Notropis telescopus	A (1), JFBM 39479, EU917440, North Fork White River, Ozark Co., MO, U.S.A.
	B (1), JFBM 41009, EU917441, Little Pigeon River, Sevierville Co., TN, U.S.A.

ASU, Arizona State University; Co., county; JFBM, James Ford Bell Museum of Natural History; UAIC, University of Alabama Ichthyological Collection; uncat., uncatalogued.

DNA sequencing

Genomic DNA was extracted using QIAamp™ tissue extraction kits (Qiagen Inc., Valencia, CA, U.S.A.) following the manufacturer's instructions. The complete mitochondrial cyt *b* gene was amplified using polymerase chain reaction (PCR). PCR reactions were performed in a total volume of 25 µl containing 1·0 µg DNA, 1·2 µM of each primer, 1×*Taq* salts, 4·0 mM $MgCl_2$, 0·4 µM dNTPs and 1·25 units of *Taq* DNA polymerase. The following thermal profile was used: initial denaturation at 94° C (3 min); 35 cycles of 94° C (10 s), 50° C (20 s), 72° C (20 s) and a final extension at 72° C (10 min) before termination of the reaction at 4° C. The heavy strand primer HA (16249) and light strand primer LA (15058) (**Schmidt *et al.*, 1998**) were used for all amplifications. Amplification products were purified using the QIAquick PCR purification kit (Qiagen Inc.) following the manufacturer's instructions. The internal primers LC (15344), HB (16002) and HD (15680) (**Schmidt *et al.*, 1998**) and amplification primers were used for sequencing. Automated sequencing was performed using Big Dye (PerkinElmer, Wellesley, MA, U.S.A.) terminator cycle sequencing on an ABI 3700 at the Advanced Genetic Analysis Center, University of Minnesota. Both strands were sequenced. Sequences were checked for accuracy of base determination and assembled using the computer programme SEQUENCHER 4.0 (Gene Codes Corp., Ann Arbor, MI, U.S.A.). All sequences were deposited in GenBank (**Table I**).

Sequence variation, phylogenetic and demographic analyses

The number of variable and potentially phylogenetically informative sites were calculated using PAUP*4.0b10 (**Swofford, 2001**). Uncorrected per cent pair-wise distances (p) were calculated using PAUP* to identify unique haplotypes. The data set was pruned to include only unique haplotypes for phylogenetic analyses.

Models of sequence evolution were chosen using MODELTEST 3.6 (**Posada & Crandall, 1998, 2001**) for the entire data set and a partitioned data set. The data set was partitioned according to codon position, and MODELTEST was run on each partition based on one tree from the parsimony analysis. Model choices were based on the Akaike information criterion (AIC) (**Posada & Buckley, 2004**). The models obtained from the partitioned data set were used as input parameters for the Bayesian analyses. The model obtained from the unpartitioned data set was subsequently used to correct pair-wise distance comparisons and to test for a molecular clock.

Bayesian analyses were carried out using MrBAYES 3.0b4 (**Ronquist & Huelsenbeck, 2003**) with the models determined by MODELTEST for each partition, random starting tree, uniform interval priors except substitution rates, which assumed a Dirichlet proposal as designated by the model. Markov chain Monte-Carlo (MCMC) was run with four chains for 2 000 000 generations. *Luxilus coccogenis* was designated as the outgroup. Trees were sampled every 100 generations; branch lengths of sampled trees were saved. The burn-in, the number of trees generated prior to the MCMC reaching stability, was determined by plotting the log-likelihood scores of sampled trees against generation time (**Huelsenbeck & Ronquist, 2001**). Trees generated after the burn-in were retained. Retained trees were used to generate a 50% majority rule consensus tree. The percentage of times each node occurs among these trees is interpreted as the posterior probability of the node (**Huelsenbeck & Ronquist, 2001**).

Four Bayesian analyses were performed to ensure that analyses were not trapped in local optima. Independent analyses were considered to have converged if their log-likelihood values approached similar mean values. The posterior probabilities for individual clades from separate analyses with similar mean log-likelihood values were compared for congruence by plotting the posterior probabilities for each node from each analysis against each other (**Huelsenbeck & Imennov, 2002**). Correlation coefficients were calculated using Excel (Microsoft Corp., Seattle, WA, U.S.A.).

Parsimony analyses were performed using the heuristic search option, 1000 random-addition sequence replicates and tree-bisection-reconnection (TBR) algorithm in PAUP* with all bases equally weighted. Trees were rooted with *L. coccogenis* and *N. telescopus*. Parsimony trees were evaluated using summary values reported by PAUP*. Support for internodes was evaluated by calculating bootstrap values (**Felsenstein, 1985**) using 1000 pseudoreplicate bootstraps with a full heuristic search, simple stepwise addition option and TBR.

Net between-group mean distances were calculated using the formula: $\delta = \delta_{xy} - (\delta_x + \delta_y)/2$, where δ_x and δ_y are the mean distances within groups x and y and δ_{xy} is the average distance between groups x and y (**Nei & Li, 1979**). Distances were calculated using the model determined by MODELTEST for the unpartitioned data set. The number of haplotypes and number of polymorphic sites were calculated using ARLEQUIN 2.000 (**Schneider *et al.*, 2000**).

The hypothesis of a constant DNA substitution rate among lineages (molecular clock) was tested using a likelihood ratio test (**Huelsenbeck & Crandall, 1997**) on the consensus tree topology for the H. amblops species group and H. amblops only. Likelihood model parameters using the substitution models determined by MODELTEST for the unpartitioned data set were input into PAUP*, and likelihood scores (−lnL) using the topology generated in the Bayesian analysis 1 were calculated with and without a molecular clock enforced. Significance was assessed using the likelihood ratio test statistic $[-2\log\Lambda = -2\ (\log L_0 - \log L_1)]$ and a χ^2 distribution (d.f. = OTU − 2) (OTU, operational taxonomic unit).

The age of extant *H. amblops* clades was estimated in a Bayesian coalescence framework using BEAST 1.4.4 (**Drummond & Rambaut, 2007**). Outgroup samples were excluded, and only *H. amblops* haplotypes were used in the analyses. MCMC chains were run for 40 000 000 generations and sampled every 1000 generations. Model parameters consisted of GTR + I +Γ substitution model with a molecular clock enforced and the coalescent Bayesian skyline plot tree prior. An unweighted pair-group method of arithmetic averages (UPGMA) was used to construct the starting tree. Two replicate analyses were performed to ensure convergence and results pooled to calculate model parameters. Because there are no external fossil or biogeographical calibrations to estimate times of divergence, dates were calculated from substitution rate estimates under a molecular clock. While there are numerous estimates of cyt *b* evolutionary rates in teleosts, lineage-specific rate heterogeneity makes it inappropriate to apply one of these rates to data from a different lineage (**Britten, 1986; Yoder & Yang, 2000**). Therefore, a uniform prior that encompassed the published range of cyt *b* substitution rates in teleosts, 0·76–2·2% Ma^{-1} (**Webb, 1998; Zardoya & Doadrio, 1999; Bowen *et al.*, 2001; Machordom & Doadrio, 2001; Perdices & Doadrio, 2001; Near & Benard, 2004**), was used to estimate the divergence times among *H. amblops* haplotypes (**Pybus *et al.*, 2003; Drummond *et al.*, 2006**). All other model parameters used default priors.

Mismatch distributions lack the ability to detect population growth under some circumstances (**Ramos-Onsins & Rozas, 2002**); therefore, tests of whether *H. amblops* population sizes have been increasing, decreasing or remain stable were performed in a Bayesian coalescence framework. The exponential growth rate (g) for each of the *H. amblops* clades under a model of exponential growth was estimated using BEAST. Clade I was excluded because of small sample size. The optimal model of sequence evolution for each clade was calculated using the AIC in MrMODELTEST 2.2 (**Nylander, 2004**). Multiple analyses were performed to ensure convergence and results pooled to calculate model parameters. As before, a uniform prior of cyt *b* substitution rates in teleosts was used; 0·76–2·2% Ma^{-1} under a molecular clock. All other model parameters used default priors. When $g > 0$, population size has been increasing; when $g < 0$, population size has been decreasing and when $g = 0$, population size has been stable. The results were interpreted conservatively by examining the distribution of g estimates from the MCMC output. If the estimate of g was positive, but the confidence intervals included zero, the population size was presumed to be stable. If the confidence intervals did not include zero, there is greater certainty in the demographic implications of the estimate. To determine if sample sizes were a factor, the R_2 (**Ramos-Onsins & Rozas, 2002**) test statistic using 10 000 coalescent simulations in DNASP 4.50.3 (**Rozas *et al.*, 2003**) was used to test whether *H. amblops* populations were stable or increasing. The R_2 test is similar in many respects to Fu's F_S (**Fu, 1997**) but is better for small sample sizes. P-values < 0·05 are considered a significant departure from the null hypothesis of constant population size.

Results

Sequence comparisons

A fragment of the cyt *b* gene, 1059 nucleotides, was sequenced for 102 individuals of *H. amblops*, five individuals of *H. amnis*, two individuals of *H. hypsinotus*, four individuals of *H. lineapunctata*, five individuals of *H. rubrifrons* and six individuals of *H. winchelli*, two individuals of *N. telescopus* and one individual of *L. coccogenis*. Positions 1060–1140 of the cyt *b* gene were trimmed from the data matrix because they were missing in most sequences. For the entire data set, there were 336 variable sites, of which 266 were potentially phylogenetically informative. For *H. amblops* only, there were 145 variable sites, of which 122 were potentially phylogenetically informative. The observed number of haplotypes, number of polymorphic sites and average within-group sequence divergence for each of the major clades identified within *H. amblops* are listed in **Table II**.

Table II. Comparison of number of haplotypes and sequence variation within *Hybopsis amblops*

	n	**Number of haplotypes**	**Number of polymorphic sites**	**Average within-group sequence divergence**
Clade I	4	2	3	0·0018
Clade II	29	11	16	0·0040
Clade III	17	13	21	0·0067
Clade IV	33	13	17	0·0027
Clade V	18	10	22	0·0087

Phylogenetic analyses

Model selection identified GTR + I +Γ as the best model for the unpartitioned data set. Model selection identified TIMef + I +Γ with the Dirichlet proposal set to 1·00, 6·55, 0·01, 0·01, 35·42 and 1·00 as the best model for partition 1, first codon position, TrN + I with the Dirichlet proposal set to 10·00, 91·90, 10·00, 10·00, 5·10 and 10·00 for partition 2, second codon position and GTR + I +Γ with a flat Dirichlet proposal for partition 3, third codon position. Bayesian analyses reached stability before 100 000 generations. Topology, branch lengths and model parameters were estimated from 19 000 of 20 000 saved trees. The relationships of *H. amblops* haplotypes were identical in all four Bayesian analyses. The phylogenetic position of *H. hypsinotus* was the only factor that varied among the analyses. In three of the analyses, it was resolved as the most basal member of the *H. amblops* species group with poor support, and in one analysis, it was included in an unresolved polytomy with the far outgroups *L. coccogenis* and *N. telescopus*. The mean −ln value for each Bayesian analysis was as follows: analysis 1 = 5084·08, analysis 2 = 5083·89, analysis 3 = 5085·28 and analysis 4 = 5084·12. The plots of the posterior probabilities supporting congruent nodes between analyses had correlation values ranging from r^2= 0·9777 to 0·9888. A high correlation of support for congruent nodes ensures convergence of analyses (**Leaché & Reeder, 2002**). Branch lengths of sampled trees for analysis 1 were calculated and used to construct the consensus tree shown in **Fig. 2**.

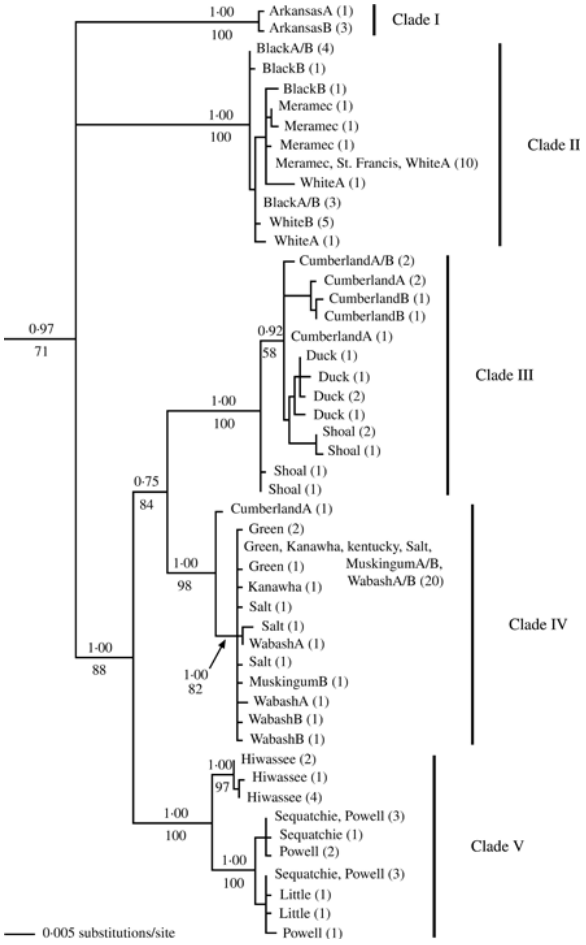

Figure 2 Topology produced in the Bayesian analysis. Consensus of 19 000 trees. Numbers above nodes indicate posterior probabilities and numbers below nodes indicate parsimony bootstrap values. Outgroups are not

shown. Numbers in parentheses indicate number of individuals with identical haplotypes. Population names correspond to **Table I**.

The parsimony analysis resulted in 3120 equally most parsimonious trees (total length = 809, excluding uninformative characters = 0·4578 and retention index = 0·8209). A strict consensus of these trees was largely consistent with the Bayesian analyses. The only observed differences in tree topology were a lack of resolution within some of the major clades identified within *H. amblops*. The strict consensus tree is not presented; bootstrap values are indicated on the Bayesian topology (**Fig. 2**).

The relationships among the putative members of the *H. amblops* species group and further outgroups are not shown. Neither the Bayesian nor parsimony analyses were able to conclusively resolve the phylogenetic position of *H. hypsinotus*. The remaining members of the species group formed a strongly supported monophyletic group (posterior probability = 100, bootstrap value = 89).

Within *H. amblops*, five major clades were resolved (**Fig. 2**): clade I, from tributaries of the Arkansas River draining the western Ozarks; clade II, from the Black, White and St Francis river drainages in the southern Ozarks and the Meramec River in the northern Ozarks; clade III, from the Cumberland and Duck Rivers and Shoal Creek from the lower Tennessee River drainage; clade IV, from the Cumberland, Green, Kanawha, Kentucky, Muskingum, Salt and Wabash River drainages and clade V, from the Hiwassee, Little, Powell and Sequatchie Rivers from the middle and upper Tennessee River drainage. Monophyly of these groups is strongly supported, but there is little resolution within them; haplotypes were widespread and not restricted to single populations or drainages. The only exception is within clade V from the middle and upper Tennessee River drainage. Haplotypes from the Hiwassee River were reciprocally monophyletic with respect to the remaining haplotypes in the clade.

Net between-group mean sequence divergences between major clades are listed in **Table III**. The greatest amount of divergence was observed between clades distributed east (clades III, IV and V) and west (clades I and II) of the Mississippi River. The null hypothesis of no rate variation among lineages was rejected for the data set containing the *H. amblops* species group plus outgroups, −2 logΛ= 135·88, $\chi^2_{63} = 128 \cdot 10$, and was not rejected for *H. amblops* only, −2 logΛ= 106·34, $\chi^2_{47} = 68 \cdot 34\alpha$= 0·001. The estimated age of clades is presented in **Table IV and Fig. 3[link]**. The large confidence intervals (95% highest posterior density) reflect the use of a single locus in these analyses. The mean age estimate for the split between haplotypes east and west of the Mississippi River is during the middle Pliocene (3·004 mya; **Table IV**). However, the confidence interval spans from the early Pliocene to the early Pleistocene (1·4792 and 4·959; **Table IV**). The mean age estimate for clades within the Ozarks west of the Mississippi River is generally older than that for the clades east of the Mississippi River, although the confidence intervals overlap. Bayesian estimates of exponential growth rate (*g*) and the model of sequence evolution used for each estimate are presented in **Table V**. The largest growth rate was observed in clade IV distributed in previously glaciated regions east of the Mississippi River. The other clades have no evidence of recent population expansion. The R_2 statistic (**Ramos-Onsins & Rozas, 2002**) was consistent with this result. The null hypothesis of a constant population size in clades II ($P < 0{\cdot}5235$), III ($P < 0{\cdot}4344$) and V ($P < 0{\cdot}8500$) could not be rejected. The R_2 statistic was significant for clade IV ($P < 0{\cdot}0364$), indicating a growing population size.

Table III. Average corrected sequence divergence between major clades of *Hybopsis amblops* identified in the Bayesian and parsimony analyses

Group comparisons	Net between-group mean distances
Clade I *v.* clade II	0·0200
Clade I *v.* clades III + IV + V	0·0239
Clade II *v.* clades III + IV + V	0·0312
Clade III *v.* clade IV	0·0163
Clade V *v.* clades III + IV	0·0171

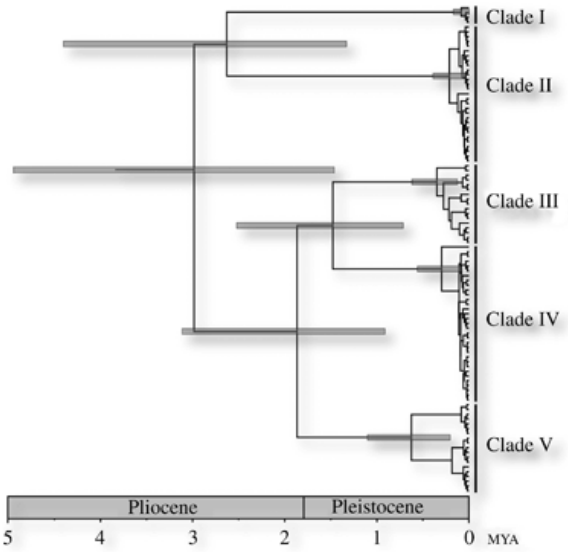

Figure 3 Bayesian consensus tree from cytochrome *b* data illustrating divergence times among major clades of *Hybopsis amblops*. Clade labels correspond to named clades in **Fig. 2**. Shaded bars indicate 95% highest posterior density of divergence times. Time is shown in million years ago (mya).

Table IV. Bayesian coalescent estimates of ages of clades of *Hybopsis amblops*. Dates are based on substitution rate estimates using a uniform prior that encompassed the published range of cytochrome *b* substitution rates in teleosts. The 95% highest posterior density (HPD) interval is given

	Mean age (mya)	95% HPD
Clade I	0·100	0·015–1·185
Clade II	0·229	0·062–0·406
Clade III	0·366	0·142–0·064
Clade IV	0·313	0·097–0·577
Clade V	0·642	0·222–1·119
Clade I + clade II	2·65	1·347–4·424
Clade III + clade IV	1·497	0·732–2·542
Clades (III + IV) + clade V	1·884	0·929–3·134
Clades (I + II) + clades (III + IV + V)	3·004	1·4792–4·959

mya, million years ago.

Table V. Bayesian estimates of the exponential growth rate (*g*), with 95% confidence intervals, for each of the *Hybopsis amblops* clades. Clade I was excluded because of small sample size. The optimal model of sequence evolution for each clade is also shown

	Model	**Exponential growth rate (*g*)**
Clade II	GTR + I	67·099 (−9·790 to 248546)
Clade III	GTR + I	7·882 (−1·532 to 20·379)
Clade IV	GTR	2424·731 (116·214–5861·644)
Clade V	HKY	0·476 (−8·638 to 9·606)

Discussion

Phylogenetic relationships

The monophyly of the *H. amblops* species group (**Mayden, 1989; Shaw *et al.*, 1995; Grose & Wiley, 2002**) was not strongly supported in the analyses. It is not clear that *H. hypsinotus* is a member of the species group. Further analyses with increased sampling are warranted to determine the status of this species. The remaining members of the species group formed a strongly supported clade. The position of *H. lineapunctata*, sister to the remaining members of the group, contradicts the hypothesis of **Grose & Wiley (2002)**. They resolved a sister group relationship between *H. lineapunctata* and *H. rubrifrons*; however, the position of this clade varied depending on the analysis performed. The basal position of *H. lineapunctata* observed in this study is more consistent with **Grose & Wiley (2002)** than with the hypotheses of **Mayden (1989) and Shaw *et al.* (1995)**, which resolved *H. lineapunctata* as a more derived member of the group. A close relationship between *H. amnis* and *H. winchelli* is consistent with the observations of **Mayden (1989), Shaw *et al.* (1995) and Grose & Wiley (2002)**. Although the putative undescribed form of *H.* sp., *cf*. *winchelli* (**Shaw *et al.*, 1995**) was not included, the paraphyly of *H. winchelli* and the close relationship of *H. amnis* and *H. winchelli* revealed that more in-depth analyses of these groups are warranted. Further discussion of the relationships among members of the *H. amblops* species group is beyond the scope of this study.

Within *H. amblops*, five strongly supported monophyletic groups were identified (**Fig. 2**). Clades I and II contained haplotypes from the Ozark Highlands. Clades III, IV and V contained haplotypes from the Eastern Highlands and previously glaciated regions of the Ohio and Wabash River drainages.

The net between-group mean distances and estimated clade ages (**Tables III and IV**) indicate that major clades within *H. amblops* diverged between the early Pliocene and the early Pleistocene (**Fig. 3**). The estimated mean age of divergence between the Ozark Highlands clades I + II and the Eastern Highlands clades III and IV + V was during the Pliocene, 3·004 mya (**Table IV**). The deep split between the Ozark and the Eastern Highlands is a pattern replicated in other clades of fishes, *e.g. H. nigricans* (**Berendzen *et al.*, 2003**) and *P. evides* (**Near *et al.*, 2001**), and is consistent with a hypothesis of a vicariant event dividing populations in drainages east and west of the Mississippi River.

Ozark highlands

Haplotypes from the Ozark Highlands comprised two monophyletic groups, clade I containing tributaries of the Arkansas River draining the western Ozarks and clade II containing drainages of the northern and southern Ozarks. The estimated mean age of these clades dates to the late Pliocene,

2·65 mya; however, the confidence interval spans from the early Pliocene to the early Pleistocene (**Table IV and Fig. 3[link]**).

The division between the Ozark and the Ouachita highlands was presumably caused by the development of the Arkansas River (**Mayden, 1985**). During the Sangamon interglacial period (0·13–0·08 mya), the Old Arkansas River captured rivers of the Plains Region including drainages of the western Ozarks (**Quinn, 1958; Mayden, 1985; Lowe & Walker, 1997**; **Fig. 4**). The split between clades I and II predates this capture. It has been hypothesized that the close relationship between tributaries of the Arkansas River draining the western Ozarks and the White River of the southern Ozarks is the result of a stream capture event (**Bretz, 1965; Kinziger *et al.*, 2007**). However, the modern Arkansas River and its associated lowlands presumably formed a barrier to gene flow in Highland fishes (**Mayden, 1985**), thus maintaining isolation of populations of *H. amblops* in the western Ozarks from the northern and southern Ozarks. This pattern has also been observed in the banded sculpin, *Cottus carolinae* (Gill) complex (**Kinziger *et al.*, 2007**).

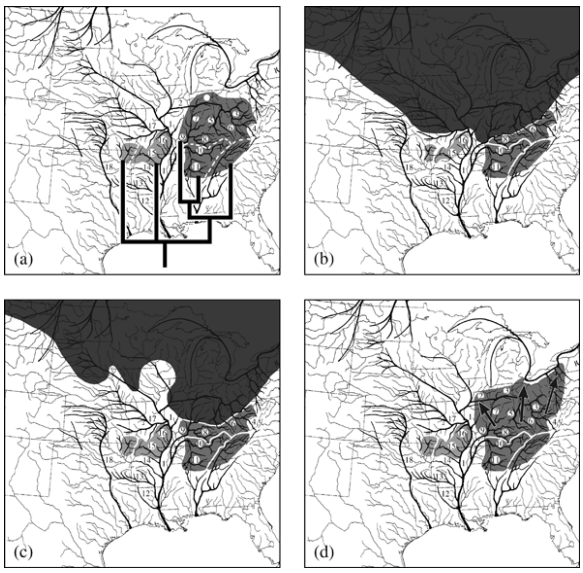

Figure 4 Reconstructed pre-glacial drainages over present drainage patterns from **Mayden (1988)**. 1, Old Mississippi River; 2, Mahomet Valley; 3, Teays River; 4, Kanawha River; 5, Old Kentucky River; 6, Old Licking River; 7, Wabash River; 8, Green River; 9, Old Ohio River; 10, Old Cumberland River; 11, Old Tennessee River; 12, Old Red River; 13, Old Ouachita River; 14, Old Arkansas River; 15, White River; 16, Old Grand-Missouri River and 17, Plains Stream. (a) Hypothetical pre-glacial distribution of *Hybopsis amblops* and the topology from **Fig. 1**. Branch lengths have no meaning. (b) Hypothetical pre-Illinoian glacial distribution of *H. amblops*. (c) Hypothetical Wisconsinan glacial distribution of *H. amblops*. (d) Post-glacial distribution of *H. amblops*. Dashed lines represent splits that occurred during the early Pleistocene. Arrows indicate demographic expansion events.

Within clade II (**Fig. 2**), there is shallow genetic divergence (0·0040 substitutions per site; **Table II**). The close relationship of haplotypes from the Meramec River in the northern Ozarks with the southern Ozarks can be explained by two alternate hypotheses; stream capture events or dispersal through the Mississippi River. Given the proximity of the headwaters of the Black and St Francis rivers with those of the Meramec River, a stream capture event is not unlikely. Alternatively, pre-Pleistocene dispersal through the Mississippi River, when it flowed in the channel of the modern Black and White rivers, is

also possible (**Fig. 4**). Both hypotheses are consistent with the faunal similarities shared between these drainages (**Cross *et al.*, 1986**), and neither hypothesis can be rejected given these data.

Clade II has an exponential growth rate (*g*) of 67; however, the lower confidence limit is less than zero, suggesting that this clade has not undergone recent expansion (**Table V**). The clade is clearly late Pleistocene in origin, haplotypes coalesce 0·229 mya (**Table IV**). West of the Mississippi River, the earliest pre-Illinoian glacial advances [**Fig. 4(b)**] extended well into Northern Missouri, whereas the younger Illinoian and most recent Wisconsinan glacial cycles [**Fig. 4(c)**] did not (**Mickelson & Colgan, 2004**). This suggests that during the pre-Illinoian glacial period, populations of *H. amblops* were displaced and contracted in the Ozark Highlands, and during the Wisconsinan glacial cycles, populations in the northern and southern Ozarks were stable (**Fig. 4**).

Eastern highlands and previously glaciated regions of the Ohio river

The patterns observed within the Eastern Highlands and previously glaciated regions of the Ohio and Wabash river drainages present a more complicated scenario than the Ozark Highlands. Haplotypes from the Eastern Highlands were recovered in a monophyletic group with three distinct clades (**Fig. 2**); clade III containing the Cumberland River and Duck River and Shoal Creek in the lower Tennessee River drainage, clade IV containing the Cumberland and Wabash rivers and the upper Ohio River drainage and clade V containing the middle and upper Tennessee River drainage. The divergence observed among these clades dates to the early Pleistocene (1·884 mya). Clade III is sister to clade IV (**Fig. 2**) with an estimated divergence of 1·497 mya (**Table III**), and clade V is sister to clades III + IV with a divergence of 1·884 mya (**Table III**). Clade V contains haplotypes restricted to the upper Tennessee River, while clades III + IV contains haplotypes from the lower Tennessee River (Duck River and Shoal Creek) as well as haplotypes from the Cumberland, Green and Ohio rivers. The upper Tennessee and lower Tennessee are thought to have been separate drainages until *c.* 1·8 mya (**Thornbury, 1965**) when the lower Tennessee captured the upper Tennessee (**Starnes & Etnier, 1986; Mayden, 1988**). Clade V has been isolated from clades III + IV for at least that long and continues to remain isolated in spite of the physical connection existing between the upper and the lower Tennessee River.

Within clade III, haplotypes are distributed in the Cumberland River and Duck River and Shoal Creek in the lower Tennessee River drainage. There is not a great deal of nucleotide variation within this clade (**Fig. 2**). The exponential growth rate is low (g = 7·811 and the lower confidence interval is <0; **Table V**), suggesting that the population is stable. Within clade IV (**Fig. 2**), haplotypes are distributed across a wide geographic area ranging from the Kanawha and Muskingum rivers in the upper Ohio to the Wabash, Green and Cumberland Rivers in the lower Ohio. It contains very shallow divergence (0·0027 substitutions per site; **Table II**). Excluding the Cumberland River, these drainages were all part of the ancient Teays River system (**Fig. 4**). East of the Mississippi River, the pre-Illinoian, Illinoian and most recent Wisconsinan glacial cycles of the Quaternary had a maximum extension as far south as modern Ohio River (**Burr & Page, 1986**). The drainages included in clade IV were most affected by glaciations, either glaciated or impounded by glaciers (**Mayden, 1988**). As ice sheets advanced southward fragmenting the Teays fauna, fishes from the middle reaches of the Teays were pushed southward and diverted into the Old Ohio River drainage (**Fig. 4**). With the retreat of the glaciers, fishes moved northward reinvading previously glaciated regions through newly established drainage systems (**Burr &**

Page, 1986). The exponential growth rate is high (g = 2424), and the lower bound is greater than zero (**Table V**). This suggests recent increase in population size. This phylogeographical pattern also observed in *P. evides* (**Near *et al.*, 2001**) and *H. nigricans* (**Berendzen *et al.*, 2003**) from the Eastern Highlands. The inclusion of one haplotype from the Cumberland River in clade III suggests that dispersion also occurred southward into the Cumberland River drainage.

In the upper Tennessee River drainage (**Fig. 2**), *H. amblops* exhibits shallow divergence (0·0087 substitutions per site; **Table II**). The low value of g indicates a stable, non-expanding population.

It is clear that neither a dispersal (**Burr & Page, 1986**) nor a vicariant (**Mayden, 1988**) hypothesis alone adequately explains the current distribution of *H. amblops*. Rather, the best explanation involves components of both early vicariant events and post-glacial range expansion. Prior to the onset of glaciation, vicariant events fragmented populations of *H. amblops* east and west of the Mississippi River leading to the split between clades I and II in the Ozark Highlands and between clades III, IV and V in the Eastern Highlands and previously glaciated regions of the Ohio River drainage. East of the Mississippi River events associated with the blocking and rerouting of Teays River system caused populations to be pushed southward into refugia of the upper Ohio River. Following the most recent Wisconsinan glacial period, populations in the upper Ohio River refugia expanded northward into previously glaciated regions and southward into the Cumberland River drainage. In the Ozarks, isolation of clades appears to be maintained by the lack of stream capture events between the upper Arkansas and the White River system and the barrier formed by the Arkansas River. East of the Mississippi River, particularly in the Tennessee River, the mechanisms responsible for maintaining isolation of clades are not clear. What was responsible for maintaining populations in the upper Tennessee River isolated from those in the lower Tennessee for 1·8 mya, while populations in the Green and Ohio were capable of dispersing over such a large area? Additional phylogeographical studies of Tennessee River fishes are needed to determine if this is a common pattern among highland fishes of the Tennessee system.

Acknowledgments

We thank M. Bush, J. Egge, M. Ghedotti, N. Gidmark, M. Johns, A. Kinziger, B. Kuhajda, B. Nagle, J. Switzer and R. Wood for assistance in the field. We also thank B. Barber and two anonymous reviewers for comments on the manuscript. Partial funding for this research was provided by the National Science Foundation (EF 0431132); the Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota; the Office of the Vice President for Research and Dean of the Graduate School, University of Minnesota and the Dayton-Wilkie Fund of the Bell Museum of Natural History, University of Minnesota.

References

Berendzen, P. B., Simons, A. M. & Wood, R. M. (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.

Berendzen, P. B., Simons, A. M., Wood, R. M., Dowling, T. E. & Secor, C. L. (2008). Recovering cryptic diversity and ancient drainage patterns in eastern North America: historical biogeography of

the *Notropis rubellus* species group (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* **46**, 721– 737.
Boschung, H. T. Jr. & Mayden, R. M. (2004). *Fishes of Alabama*. Washington, DC: Smithsonian Books.
Bowen, B. W., Bass, A. L., Rocha, L. A., Grant, W. S. & Robertson, D. R. (2001). Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale. *Evolution* **55**, 1029– 1039.
Bretz, J. H. (1965). *Geomorphic history of the Ozarks of Missouri*. Rolla, MO: Missouri Geological Survey and Water Resources.
Britten, R. J. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**, 1393– 1398.
Burr, B. M. & Page, L. M. (1986). Zoogeography of fishes of the lower Ohio–upper Mississippi Basin. In *The Zoogeography of North American Fishes* (C. H. Hocutt & E. O Wiley., eds), pp. 287– 324. New York, NY: Wiley Interscience.
Cross, F. B., Mayden, R. L. & Stewart, J. D. (1986). Fishes in the western Mississippi Basin (Missouri, Arkansas and Red rivers). In *The Zoogeography of North American Fishes* (C. H. Hocutt & E. O Wiley., eds), pp. 363– 412. New York, NY: Wiley Interscience.
Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.
Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88.
Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783– 791.
Fisk, H. N. (1944). *Geological Investigation of the Alluvial Valley of the Lower Mississippi River*. Vicksburg, MS: Mississippi River Commission.
Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915– 925.
Gibbard, P. & Van Klofschoten, T. (2004). The Pleistocene and Holocene epochs. In *A Geologic Time Scale 2004* (F. M. Gradstein, J. G. Ogg & A. G Smith., eds), pp. 441– 452. Cambridge: Cambridge University Press.
Grose, M. J. & Wiley, E. O. (2002). Phylogenetic relationships of the *Hybopsis amblops* species group (Teleostei: Cyprinidae). *Copeia* **2002**, 1092– 1097.
Huelsenbeck, J. P. & Crandall, K. A. (1997). Phylogenetic estimation and hypothesis testing using maximum likelihood. *Annual Review Ecology and Systematics* **28**, 437– 466.
Huelsenbeck, J. P. & Imennov, N. S. (2002). Geographic origin of human mitochondrial DNA: accommodating phylogenetic uncertainty and model comparison. *Systematic Biology* **51**, 155– 165.
Huelsenbeck, J. P. & Ronquist, F. R. (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 154– 755.
Jenkins, R. E. & Burkhead, N. M. (1994). *Freshwater Fishes of Virginia*. Bethesda, MD: American Fisheries Society.
Kinziger, A. P., Goodman, D. H. & Studebaker, R. S. (2007). Mitochondrial DNA variation in the Ozark Highland members of the banded sculpin *Cottus carolinae* complex. *Transactions of the American Fisheries Society* **136**, 1742– 1749.
Leaché, A. D. & Reeder, T. W. (2002). Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of parsimony, likelihood, and Bayesian approaches. *Systematic Biology* **51**, 44– 68.

Lourens, L., Hilgen, F., Shackleton, N. J., Laskar, J. & Wilson, D. (2004). The Neogene period. In *A Geologic Time Scale 2004* (F. M. Gradstein, J. G. Ogg & A. G Smith., eds), pp. 409– 440. Cambridge: Cambridge University Press.

Lowe, J. J. & Walker, M. J. C. (1997). *Reconstructing Quaternary Environments*, 2nd edn. Essex: Longman.

Machordom, A. & Doadrio, I. (2001). Evidence of a Cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (*Luciobarbus*, Cyprinidae). *Molecular Phylogenetics and Evolution* **18**, 252– 263.

Mayden, R. L. (1985). Biogeography of Ouachita Highland fishes. *The Southwestern Naturalist* **30**, 195– 211.

Mayden, R. L. (1988). Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Systematic Zoology* **37**, 329– 355.

Mayden, R. L. (1989). Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: cypriniformes). *University of Kansas Museum of Natural History Miscellaneous Publications* **80**, 1– 189.

Mickelson, D. M. & Colgan, P. M. (2004). The southern Laruentide Ice Sheet. In *The Quaternary Period in the United States* (A. R. Gillespie, S. C. Porter & B. F Atwater., eds), pp. 1– 16. Amsterdam: Elsevier.

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

Near, T. J., Page, L. M. & Mayden, R. L. (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., Kassler, T. W., Koppelman, J. B., Dillman, C. B. & Philipp, D. P. (2003). Speciation in North American black basses, *Micropterus* (Actinopterygii: Centrarchidae). *Evolution* **57**, 1610– 1621.

Nei, M. & Li, W. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA* **76**, 5269– 5273.

Nylander, J. A. A. (2004). *MrModeltest v2*. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University.

Perdices, A. & Doadrio, I. (2001). The molecular systematics and biogeography of the European cobitids based on mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* **19**, 468– 478.

Pflieger, W. L. (1971). A distributional study of Missouri fishes. *University of Kansas Museum of Natural History Publication Series* **20**, 225– 570.

Posada, D. & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**, 793– 808.

Posada, D. & Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817– 818.

Posada, D. & Crandall, K. A. (2001). Selecting the best fit model of nucleotide substitution. *Systematic Biology* **50**, 580– 601.

Pybus, O. G., Drummond, A. J., Nakano, T., Robertson, B. H. & Rambaut, A. (2003). The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. *Molecular Biology and Evolution* **20**, 381– 387.

Quinn, J. H. (1958). Plateau surfaces of the Ozarks. *Proceedings of the Arkansas Academy of Science* **11**, 36– 43.

Ramos-Onsins, S. E. & Rozas, J. (2002). Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19**, 2092– 2100.

Robison, H. W. (1986). Zoogeographic implications of the Mississippi River Basin. In *The Zoogeography of North American Fishes* (C. H. Hocutt & E. O Wiley., eds), pp. 267– 286. New York, NY: Wiley Interscience.

Ronquist, F. R. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572– 1574.

Rozas, J., Sánchez-Del Barrio, J. C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496– 2497.

Schmidt, T. R., Bielawski, J. P. & Gold, J. R. (1998). Molecular phylogenetics and evolution of the cytochrome *b* gene in the cyprinid genus *Lythrurus* (Actinopterygii: Cypriniformes). *Copeia* **1998**, 14– 22.

Schneider, S., Roessle, D. & Excoffier, L. (2000). *Arelequin ver. 2.000: A Software for Population Genetics Data Analysis*. Geneva: Genetic and Biometry Laboratory, University of Geneva.

Shaw, K. A., Wiley, E. O. & Titus, T. A. (1995). Phylogenetic relationships among members of the *Hybopsis amblops* species group (Teleostei: Cprinidae). *Occasional Papers of the Museum of Natural History, University of Kansas* **172**, 1– 28.

Simons, A. M. (2004). Phylogenetic relationships in the genus *Erimystax* (Actinopterygii: Cyprinidae) based on the cytochrome *b* gene. *Copeia* **2004**, 351– 356.

Starnes, W. C. & Etnier, D. A. (1986). Drainage evolution and fish biogeography of the Tennessee and Cumberland rivers drainage realm. In *The Zoogeography of North American Fishes* (C. H. Hocutt & E. O Wiley., eds), pp. 325– 361. New York, NY: Wiley Interscience.

Swofford, D. L. (2001). *PAUP*: Phylogenetic Analysis Using Parsimony, Version 4.0b10*. Sunderland, MA: Sinauer Associates.

Thornbury, W. D. (1965). *Regional Geomorphology of the United States*. New York, NY: John Wiley and Sons.

Webb, S. A. (1998). A phylogenetic analysis of the Goodeidae (Teleostei: Cyprinodontiformes). PhD Thesis, University of Michigan, Ann Arbor, MI.

Wiley, E. O. & Mayden, R. L. (1985). Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. *Annals of the Missouri Botanical Garden* **72**, 596– 635.

Yoder, A. D. & Yang, Z. H. (2000). Estimation of primate speciation dates using local molecular clocks. *Molecular Biology and Evolution* **17**, 1081– 1090.

Zardoya, R. & Doadrio, I. (1999). Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *Journal of Molecular Evolution* **49**, 227– 237.