

EFFECTS OF CLIMATIC AND GEOLOGICAL PROCESSES DURING THE PLEISTOCENE ON THE EVOLUTIONARY HISTORY OF THE NORTHERN CAVEFISH, *AMBLYOPSIS SPELAEA* (TELEOSTEI: AMBLYOPSIDAE)

Matthew L. Niemiller,^{1,2,3} James R. McCandless,⁴ R. Graham Reynolds,^{2,5} James Caddle,² Thomas J. Near,¹ Christopher R. Tillquist,⁶ William D. Pearson,⁷ and Benjamin M. Fitzpatrick²

¹Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut 06520

²Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee 37996

³E-mail: matthew.niemiller@yale.edu

⁴Department of Biology, Methodist University, Fayetteville, North Carolina 28311

⁵Department of Biology, University of Massachusetts Boston, Boston, Massachusetts 02125

⁶Department of Anthropology, University of Louisville, Louisville, Kentucky 40292

⁷Department of Biology, University of Louisville, Louisville, Kentucky 40292

Received August 23, 2012

Accepted October 28, 2012

Data Archived: doi:10.5061/dryad.25kd9

Climatic and geological processes associated with glaciation cycles during the Pleistocene have been implicated in influencing patterns of genetic variation and promoting speciation of temperate flora and fauna. However, determining the factors promoting divergence and speciation is often difficult in many groups because of our limited understanding of potential vicariant barriers and connectivity between populations. Pleistocene glacial cycles are thought to have significantly influenced the distribution and diversity of subterranean invertebrates; however, impacts on subterranean aquatic vertebrates are less clear. We employed several hypothesis-driven approaches to assess the impacts of Pleistocene climatic and geological changes on the Northern Cavefish, *Amblyopsis spelaea*, whose current distribution occurs near the southern extent of glacial advances in North America. Our results show that the modern Ohio River has been a significant barrier to dispersal and is correlated with patterns of genetic divergence. We infer that populations were isolated in two refugia located north and south of the Ohio River during the most recent two glacial cycles with evidence of demographic expansion in the northern isolate. Finally, we conclude that climatic and geological processes have resulted in the formation of cryptic forms and advocate recognition of two distinct phylogenetic lineages currently recognized as *A. spelaea*.

KEY WORDS: Cryptic diversity, historical demography, population genetics, refugia, subterranean.



Pleistocene glacial ice sheets have advanced several times southward from the Canadian Arctic during the last 2.7 million years. These glacial advances and associated climatic fluctuations have shaped the distributions and patterns of genetic variation of many temperate flora and fauna (Pielou 1991; Hewitt 2000, 2004) and have played a key role in promoting speciation in several lineages (Johnson and Cicero 2004; Ribera and Vogler 2004). In addition to isolating populations in multiple refugia south of the ice sheets, glacial cycles during the Pleistocene dramatically altered hydrological flows and even formed new river systems in North America that may be important geological barriers to gene flow in amphibians (Kozak et al. 2006; Lemmon et al. 2007). These cycles resulted in genetic divergence and potentially facilitated speciation in several other groups (Carstens and Knowles 2007), including several lineages of ray-finned fishes (Mayden 1988; Strange and Burr 1997; Berendzen et al. 2003).

Pleistocene glacial advances also had a significant impact on the distribution and diversity of subterranean fauna in the Northern Hemisphere (Porter et al. 2007; Folquier et al. 2008; Culver and Pipan 2009, 2010). Troglabites (obligate subterranean fauna) are largely absent from formerly glaciated areas and it is thought that most fauna located north of the southern limit of Pleistocene glaciers were extirpated as a result of dramatic changes in climatic and hydrological conditions, including declines in groundwater temperatures, altered hydrological regimes, and lower allochthonous organic inputs (Peck and Christiansen 1990). However, a few invertebrate species (primarily aquatic) are found in caves of formerly glaciated regions and their existence reflects either post-glacial colonization from ice-free regions or survival in subglacial refugia (Holsinger 1980; Lewis 1983; Peck and Christiansen 1990; Kristjansson and Svavarsson 2007; Kornobis et al. 2010).

During glacial maxima of the Pleistocene, most of North America was completely covered by glaciers reaching as far south as the modern Ohio River in southern Indiana and Ohio in the eastern United States. However, much of the narrow arc of karst in the Crawford-Mammoth Cave Uplands and Mitchell Plain in south-central Indiana remained free of ice (Fig. 1), and this region may have served as a periglacial refugium for many small, invertebrate troglabites and the source of migrants for subterranean habitats in adjacent, previously glaciated regions (Lewis 1983). Alternatively, periglacial conditions in south-central Indiana may have been too harsh for many troglabitic species and current subterranean biodiversity in this region would then reflect post-glacial colonization from karst areas to the south (Barr 1960; Niemiller and Poulson 2010). However, many species of *Pseudanophthalmus* caves beetles appear to have persisted in periglacial refugia in southern Indiana (Barr 1960). Although some smaller invertebrates likely persisted throughout the Pleistocene in periglacial (e.g., *Pseudanophthalmus* cave beetles), or in some cases even

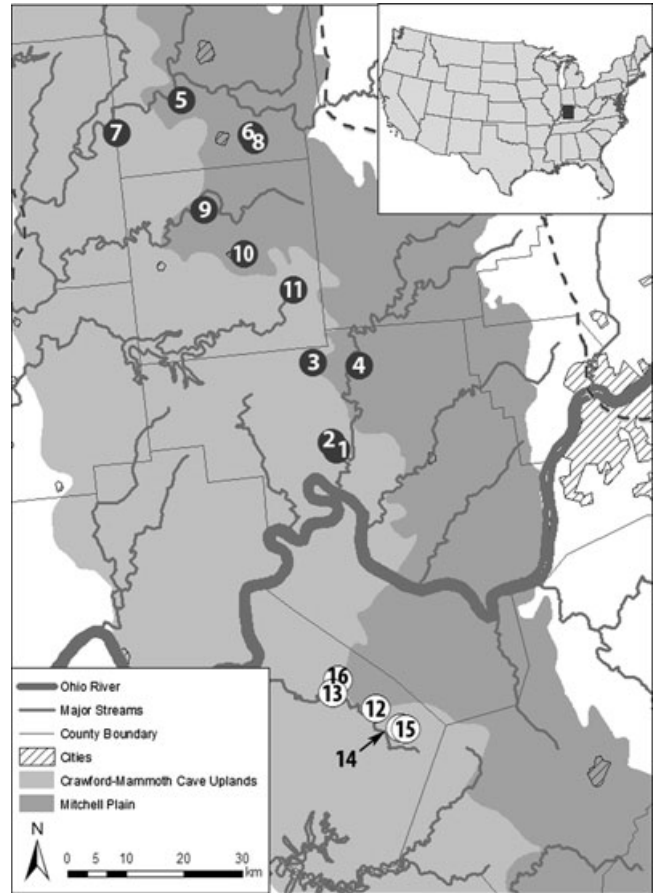


Figure 1. Map showing the distribution and sampling localities of *Amblyopsis spelaea* with respect to the Ohio River (bold dark grey line). Numbered localities correspond to populations listed in Table 1. Black circles represent populations belonging to the northern group whereas white circles represent populations belonging to the southern group. Populations exist within the Crawford-Mammoth Cave Uplands (lighter grey) and Mitchell Plain (darker grey) ecoregions in an area that remained unglaciated throughout the Pleistocene. The southern extent of Pleistocene glaciations is indicated by a dashed black line.

subglacial (e.g., groundwater amphipods in Iceland) refugia, similar evidence for larger troglabites that occupy higher trophic levels in subterranean ecosystems, such as crayfishes and cavefishes, is largely lacking.

Drainage patterns in this region were very different before the Pleistocene when present-day tributaries of the Ohio River were headwater drainages in eastern Kentucky to the ancient Teays River (Gray 1991; Fig. 1 in Teller and Goldthwait 1991). The modern course of the Ohio River formed approximately 0.8 mya (Gray 1991) associated with multiple episodes of glacial meltwater discharge and upstream drainage capture by the Old Ohio River following the overflow and draining of Lake Tigh (Teller and Goldthwait 1991). Although initially shallow, the modern Ohio River is now deeply entrenched through cave-bearing strata

throughout portions of its course through the Crawford-Mammoth Cave Uplands and Mitchell Plain. The river is now 75 m below the bottom of the Teays-aged valleys and has completely bisected the cave-bearing St. Genevieve and St. Louis limestones in parts of its course (Teller and Goldthwait 1991) likely isolating cave faunas on either side of the river. Even in areas where cave-bearing strata have not been completely bisected, it is highly likely that caverns or solution channel networks below the Ohio River were quickly filled by sediment or glacial outwash, as water velocities continually decreased as the Ohio River became larger with less of a gradient (McCandless 2005). If the Ohio River has been a barrier, then we should observe genetic divergence between populations north and south of the river correlated with timing of formation of the modern course of the Ohio River. However, several troglotic species are distributed both north and south of the modern Ohio River (Hobbs et al. 1977; Niemiller and Poulson 2010) and it is unclear whether the river truly is a significant obstacle to gene flow or if divergence predates the formation of the Ohio River and the Pleistocene altogether. In addition, patterns of genetic variation and divergence may be shaped by other factors, such as geologic factors or climatic events unrelated to the Pleistocene (Juan and Emerson 2010; Juan et al. 2010, and references therein). For example, several phylogeographic studies of subterranean fauna have found that diversification often is associated with climatic or geologic events that predate the Pleistocene (e.g., Ribera et al. 2010; Faille et al. 2010).

The Northern Cavefish (*Amblyopsis spelaea*) is a teleost fish (family Amblyopsidae) that occurs in subterranean waters throughout the narrow arc of karst of the Crawford-Mammoth Cave Uplands and Mitchell Plain from south-central Indiana to the Mammoth Cave area of central Kentucky (Niemiller and Poulson 2010). The distribution of *A. spelaea* north of the modern Ohio River in Indiana lies within the area that remained ice free throughout the Pleistocene but occurs as close as 16 km from the glacial maxima during the Illinoian glaciation (300–130k years before present). The divergence of *A. spelaea* from its closest surface-dwelling relatives (*Forbesichthys agassizii* and *F. papilliferus*) occurred 4.0 to 7.9 mya (Niemiller et al. 2013), suggesting a much longer history of subterranean isolation than the well-known Mexican cave tetra, *Astyanax mexicanus*, which appears to have several more recent and independently derived cave populations (Porter et al. 2007). However, patterns of genetic divergence among *A. spelaea* populations have not yet been investigated. Moreover, several recent studies have documented significant cryptic diversity in aquatic subterranean fauna often associated with hydrological drainages (Verovnik et al. 2003; Finston et al. 2007; Niemiller et al. 2012). Because its distribution includes at least seven hydrological basins, *A. spelaea* might comprise several morphologically cryptic lineages, as has been documented in other amblyopsid cavefishes (Niemiller

et al. 2012), in which divergence and genetic structure have been shaped by past geological and climatic processes. Limited data on rudimentary eye size and pigmentation suggest that the Ohio River might indeed be a barrier (Poulson 1960).

We hypothesize that patterns of genetic variation in *A. spelaea* have been influenced by both geological and climatic processes that occurred during the Pleistocene. We take an integrative hypothesis-testing approach to assess the roles of vicariance and dispersal and to examine if cryptic diversity is consistent with the evolutionary history of *A. spelaea*. First, we assess if the formation of the Ohio River is correlated with the partitioning of genetic variation and divergence within *A. spelaea*. Along this vein, we also investigate whether genetic variation may be associated with other factors, including hydrological boundaries of surface subbasins and geologically defined ecoregions. Second, we determine if climatic fluctuations significantly influenced genetic diversity by assessing whether the current distribution of *A. spelaea* reflects persistence throughout the Pleistocene *in situ* or isolation in periglacial refugia followed by colonization northward. We employ a multilocus dataset to address these questions by correlating fossil-calibrated, coalescent-based divergence time estimates among lineages to the timing of geological and climatic events and by examining and comparing levels of genetic variation throughout the distribution of *A. spelaea*.

Materials and Methods

SPECIMEN SAMPLING

We collected specimens or tissue samples (fin clips) from 72 individuals of 16 populations across the distribution of *A. spelaea* (Fig. 1) in Indiana and Kentucky (Table 1). The distribution of *A. spelaea* is largely continuous; however, the species is found in two distinct clusters of populations south of the Ohio River: one in the Mammoth Cave region in south-central Kentucky and the other in Breckinridge Co., Kentucky, primarily within the Sinking Creek drainage, which flows into the Ohio River. These clusters are likely separated from each other by the Rough Creek Fault Zone (also known as the Hart County Ridge) from Hart County to Webster and Union counties in Kentucky. Few localities are known outside these two regions and recent surveys of caves in the vicinity of the Rough Creek Fault Zone failed to yield cavefish. Consequently, our sampling south of the Ohio River focused on the core set of populations in Breckinridge County closest to the Ohio River relevant to the questions addressed in this study. We also included in our analyses samples or sequences for all other amblyopsid cavefishes as outgroups, including *Chologaster cornuta*, *F. agassizii*, *F. papilliferus*, *Speoplatyrhinus poulsoni*, *Troglichthys rosae*, *Typhlichthys eigenmanni*, and *T. subterraneus* for divergence time analyses. Samples and sequences of other

Table 1. Locality information, including cave, county, state, sample, hydrological basin and subbasin, and ecoregion, and haplotype numbers observed at each locality for 16 populations of *Amblyopsis spelaea*. Eleven populations ($n = 36$) were sampled north (N) and five populations ($n = 36$) south (S) of the Ohio River.

No.	Group	Locality	Abbrev.	County	State	n	Basin: Subbasin	Ecoregion	$nd2$	$s7$	$rag1$	tbr	ρ
1	N	BB Hole	BBH	Crawford	IN	2	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	1	1	1	1	1
2	N	Eric's River Cave	ERI	Crawford	IN	4	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	3, 6	1	1	1	1
3	N	Marengo New Discovery	MND	Crawford	IN	3	Lower Ohio: Blue-Sinking	Mitchell Plain	1	1	1	1	1
4	N	Black Medusa Cave	BLM	Harrison	IN	5	Lower Ohio: Blue-Sinking	Mitchell Plain	1	1	1	1	1
5	N	Blue Springs Caverns	BLS	Lawrence	IN	11	Wabash: Lower East Fork White	Mitchell Plain	1	1	1	1	1
6	N	Donaldson Cave	DON	Lawrence	IN	3	Wabash: Lower East Fork White	Mitchell Plain	2	1	1	1	1
7	N	Henshaw Bend Cave	HSB	Lawrence	IN	1	Wabash: Lower East Fork White	Crawford-Mammoth Cave Uplands	5	1	1	1	1
8	N	Upper Twin Cave	UPT	Lawrence	IN	1	Wabash: Lower East Fork White	Mitchell Plain	7	1	1	1	1
9	N	Elrod Cave	ELR	Orange	IN	1	Wabash: Lower East Fork White	Mitchell Plain	1	2	1	1	1
10	N	Murray Spring Cave	MSP	Orange	IN	1	Wabash: Lower East Fork White	Mitchell Plain	1	1	2	1	1
11	N	Spring's Spring Cave	SPR	Orange	IN	4	Wabash: Patoka	Crawford-Mammoth Cave Uplands	1, 4	1	1, 3	1	1
12	S	Bandy Cave	BND	Breckinridge	KY	3	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	10	2	1	2	2
13	S	Penitentiary Cave	PEN	Breckinridge	KY	11	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	8	2	1, 4	1	2
14	S	Rimstone Cave	RIM	Breckinridge	KY	10	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	8	2	1	1	2
15	S	Under the Road Cave	UTR	Breckinridge	KY	4	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	8, 11	2	1, 4	1	2
16	S	Webster Cave	WEB	Breckinridge	KY	8	Lower Ohio: Blue-Sinking	Crawford-Mammoth Cave Uplands	8, 9	2	1	1, 3	2

related species that represent major lineages of percopsiform fishes (*Aphredoderus gibbosus*, *A. sayanus*, *Percopsis omiscomaycus*, and *P. transmontana*) were also included because the fossil constraints employed to estimate divergence dates fall outside the amblyopsid clade (Niemiller et al. 2013).

DNA EXTRACTION, POLYMERASE CHAIN REACTION (PCR), AND SEQUENCING

Genomic DNA was extracted using Qiagen DNEasy kits (Qiagen Inc.) or by standard phenol–chloroform methods. We used PCR to

amplify all or parts of one mitochondrial and four nuclear genes (Table S1) following previous published protocols (Niemiller et al. 2012, 2013), including mitochondrial NADH dehydrogenase 2 ($nd2$) and nuclear intron 1 of ribosomal protein $s7$ ($s7$), exon 3 of recombination activating gene 1 ($rag1$), rhodopsin (ρ), and T-box brain 1 ($tbr1$). Purified PCR products were sequenced at the Molecular Biology Resource Facility, Division of Biology, University of Tennessee. Our dataset was also supplemented with available sequences previously accessioned on GenBank for related studies (Niemiller et al. 2012, 2013).

Sequences were aligned and edited in SEQUENCHER v4.5 (Gene Codes) with resulting contigs aligned in MACCLADE v4.07 (Maddison and Maddison 2005). Alleles of nuclear sequences that contained more than one heterozygous site were inferred in a Bayesian framework with PHASE v2.1 (Stephens et al. 2001; Stephens and Scheet 2005) using 10,000 iterations and a burn-in of 10,000 generations. To check for consistency of results, haplotype frequency estimates and goodness-of-fit measures from three independent runs were compared. No sequences were excluded from analyses. We accessioned unique DNA sequences generated during this study into GenBank (JX977875–JX979154).

ESTIMATION OF GENE TREES AND HAPLOTYPE NETWORKS

Gene trees for each locus were constructed using partitioned Bayesian analyses, and posterior probabilities were estimated with the MCMC algorithm implemented in MrBAYES 3.1 (Ronquist and Huelsenbeck 2003). All loci are protein coding and were partitioned by codon with the exception of the first intron of ribosomal protein *s7*. The best-fit models of molecular evolution for each partition were selected using the Akaike's information criterion (AIC) implemented in MODELTEST v3.7 (Posada and Crandall 1998). Conditions for the MCMC followed Niemiller et al. (2012). Samples from the stationary distribution of trees were used to generate 50% majority-rule consensus trees for each locus. We also constructed unrooted statistical parsimony haplotype networks for all loci in TCS v1.21 (Clement et al. 2000) to visualize the number of mutations between groups of populations.

TESTS FOR BARRIERS TO DISPERSAL

To examine whether the Ohio River has been a significant barrier to dispersal among populations, we conducted a partial Mantel test on each locus (Mantel 1967; Smouse et al. 1986) to test for a correlation between genetic distance and position relative to the Ohio River while controlling for geographic distance. If the river is a barrier, then genetic distances between pairs of populations on the same side of the river should be lower than genetic distances between pairs of populations spanning the river. We first computed a matrix of pairwise uncorrected genetic distances for each, whereas a matrix of binary variables was calculated where the position of two populations relative to the Ohio River was coded as either occurring on the same side or on opposite sides of the river. Then we calculated a matrix of geographic distances as the great-circle distance between a pair of populations. All partial Mantel tests were calculated using ZT v1.1 (Bonnet and Van de Peer 2002) with 100,000 permutations.

As an additional test of the Ohio River barrier hypothesis, we grouped populations by regions north and south of the Ohio River and used hierarchical analyses of molecular variance

(AMOVA; Excoffier et al. 1992) implemented in ARLEQUIN v3.5 (Excoffier and Lischer 2010) using uncorrected pairwise distances and 10,000 permutations. Hierarchical AMOVA partitions the total genetic variance into covariance components due to differences among *a priori* groups, among populations within groups, and within populations. As an alternative to the Ohio River barrier, we also examined the partitioning of genetic variation by hydrological subbasins and ecoregions (Table 1). Interconnectivity of drainage basins and ecoregions shape genetic structure in other aquatic, subterranean taxa in the Interior Highlands of eastern North America (Niemiller et al. 2008, 2012) and may also have had an effect on patterns of genetic structure in *A. spelaea*. We conducted both partial Mantel tests and AMOVAs as outlined above but grouping populations by hydrological subbasins and ecoregions.

DIVERGENCE TIME ESTIMATES

If the formation of the Ohio River has been a significant barrier facilitating divergence in *A. spelaea*, then the date of this geological event (ca. 0.8 mya; Gray 1991) should fall within the confidence intervals for the estimated divergence time between populations north and south of the river. To investigate timing of diversification, we estimated divergence times using the Bayesian, coalescence-based program *BEAST v1.6.1 (Drummond and Rambaut 2007; Heled and Drummond 2010) in a species tree framework that uses multilocus data to jointly estimate multiple gene trees embedded in a shared species tree under the multispecies coalescent. We included five samples from each *Amblyopsis* group (north and south of the Ohio River) and two from all other percopsiform species. Sequence data were partitioned by locus and by codon position for protein-coding loci. Partition-specific models of nucleotide substitution (Table S1) were implemented, all parameters were unlinked across loci (not across data partition), and an uncorrelated lognormal (UCLN) model of rate variation was assumed for each partition. A Yule process speciation prior was used for the branching rates.

Because no amblyopsid fossils exist, we used two fossil calibration age prior distributions from non-amblyopsid fossil taxa. †*Tricophanes foliarum* Cope (1872) is known from the Oligocene and shares common ancestry with *Aphredoderus* (Rosen 1962; Rosen and Patterson 1969). The age of the node containing the Aphredoderidae and Amblyopsidae was calibrated using the age of this fossil. We chose a lognormal distribution such that the minimum possible sampled age corresponded to 33.9 Ma. †*Lateopisciculus turrifumosus* (Murray and Wilson 1996) is known from the Paleocene and shares common ancestry with *Percopsis* (Murray and Wilson 1996). We calibrated the most recent common ancestor of Percopsiformes using the age of this fossil, choosing a lognormal distribution such that the minimum possible sampled age corresponded to 58.7 Ma. Following McCormack

et al. (2011), we hand-edited the XML file to incorporate fossil priors on the species tree. We conducted three independent MCMC runs for 100 million generations for each analysis, sampling every 2000 generations. All runs were examined in TRACER v1.5 to monitor convergence and likelihood stationarity and verify that an effective sample size (ESS) exceeded 200 for all parameters being estimated. A conservative burn-in of 40 million generations was excluded from each run. The tree and log files were combined using LOGCOMBINER (v1.6.1, distributed as part of the BEAST package). The maximum credibility tree with mean node heights was recovered in TREEANNOTATOR (v1.6.1, distributed as part of the BEAST package).

TESTS FOR POPULATION EXPANSION

If periglacial conditions were too harsh, even in subterranean habitats, then the present-day distribution of *A. spelaea* would reflect range expansion from areas further south. If this were the case, then we predict that populations north of the Ohio River would have lower genetic variation than populations south of the river and there would be a signal of recent population growth north of the Ohio River over the last 10 to 20 thousand years. We calculated measures of genetic diversity, including the number of unique haplotypes (K), the number of segregating sites (S), and nucleotide diversity (π), in ARLEQUIN for each group of populations north and south of the Ohio River and overall for each locus.

To test for departures from neutrality or constant population size, we calculated the summary statistics F_s (Fu 1997), Tajima's D (Tajima 1989), and R_2 (Ramos-Onsins and Rozas 2002). Significant negative values of F_s and Tajima's D and small positive values of R_2 indicate population growth. Neutrality tests on each locus were performed in DNASAM (Eckert et al. 2010) for each group of populations (north and south) and all populations combined. Significance was determined by 10,000 permutations.

We also reconstructed the demographic history of *A. spelaea* using GMRF skyride plots (Minin et al. 2008) implemented in BEAST. GMRF skyride plots are a nonparametric approach that incorporates the waiting time between coalescent events in a gene tree to estimate changes in effective population size over time. We constructed GMRF skyride plots for each group of populations north and south of the Ohio River using the *nd2* dataset only because of low variability for nuclear loci. The estimated posterior rate of molecular evolution for the *nd2* locus was determined from the *BEAST divergence time analysis. We ran the analysis twice for 20 million generations each sampling every 2000 generations. All runs were visualized in TRACER and to verify that ESS values exceeded 200 for all parameters estimated. A conservative burn-in of 5 million generations was excluded from each run. Changes in effective population size over time were deemed significant if the upper and lower 95% confidence intervals at the root of the plot did not overlap those at the tips (Eytan and Hellberg 2010).

DETECTING CRYPTIC DIVERSITY

Several methodologies have recently been developed to delimit species and uncover cryptic diversity using molecular data (Knowles and Carstens 2007; O'Meara 2010; Yang and Rannala 2010; Ence and Carstens 2011). A consensus has yet to be reached as to which approach is most appropriate, however there is a general view that the use of multilocus datasets is warranted for accurate species delimitation (O'Meara 2010; Yang and Rannala 2010; Kubatko et al. 2011; Niemiller et al. 2012). Here we employ recently developed approaches to species delimitation that make use of multilocus data operating under the logic that consistent delimitation of sets of populations as distinct lineages across methods provides stronger support for species recognition than the results of a single approach alone.

First, we used the nonparametric heuristic method described in O'Meara (2010) to jointly delimit species and estimate the species tree using a multilocus dataset in the program BROWNIE v2.1 (O'Meara et al. 2006). This approach apports individuals into putative species by attempting to minimize excess intraspecific structure while minimizing gene tree conflict among species. Because many individuals shared identical alleles we included seven individuals of each group in *Amblyopsis* and *Forbesichthys* corresponding to the maximum number of alleles observed for a locus. Heuristic searches were conducted with the number of random starting species trees (NReps) set to 100, all possible taxon reassignments on leaf splits were explored (Subsample = 1), the minimum number of species (MinNumSpp) was set to 2, the maximum number of species (MaxNumSpp) was set to 6, and the minimum number of samples per species (MinSamp) was set to 2. The 50% majority-rule consensus gene trees generated from Bayesian analyses were used as input trees. We conducted 5000 independent BROWNIE runs on the BulldogK cluster at Yale University.

As an additional measure to delimit species without *a priori* group assignments, we investigated genetic structure in *Amblyopsis spelaea* using the MCMC clustering algorithm STRUCTURE 2.3.3 (Pritchard et al. 2000). Haplotypes for each locus were treated as alleles (e.g., Eytan and Hellberg 2010) and 10 independent runs were conducted for each value of $K = 1$ to 10, with 100,000 generations of burn-in and 1 million post-burn-in replicates using the admixture model. Values of K were compared by the ΔK method (Evanno et al. 2005) to infer the best estimate of K .

Although species discovery methods are advantageous in systems that have not been well studied to develop taxonomic hypotheses within defining groups *a priori*, such as many subterranean organisms (Niemiller et al. 2012), it does not incorporate other sources of existing data available for more well-studied groups. Species validation methods (Cummings et al. 2008; Yang and Rannala 2010; Ence and Carstens 2011) quantify support for *a priori* groupings of samples that are based on other lines

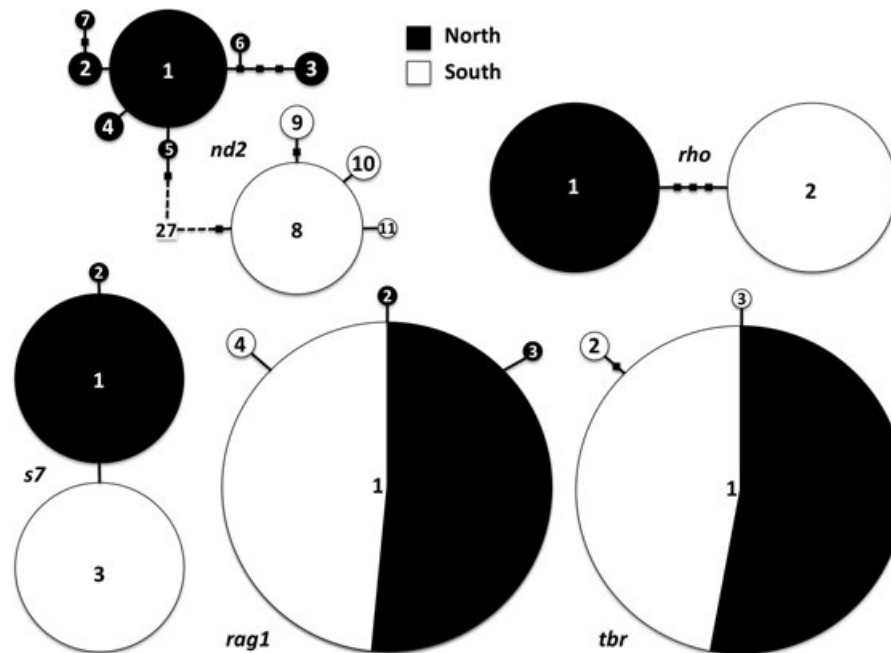


Figure 2. Haplotype networks for *Amblyopsis* for each locus. Circle color indicates the group (black for populations north of the Ohio River and white for south of the Ohio River) and size is proportional to the number of individuals sharing that haplotype. Small black squares on branches are inferred mutations not sampled. Numbers on circles correspond to distinct haplotypes listed in Table 1. Twenty-seven mutations separate the *nd2* networks north and south of the Ohio River.

of evidence, such as morphological, geographical, and behavioral data. Geological and geographic evidence suggest that the Ohio River is a vicariant barrier for subterranean fauna and taxa distributed on opposite sides of the river, including *Amblyopsis spelaea*, might constitute independent lineages. To test this hypothesis, we employed two species validation methods: Bayesian species delimitation (Yang and Rannala 2010) and genealogical tests of taxonomic distinctiveness (Cummings et al. 2008).

We conducted Bayesian species delimitation (Yang and Rannala 2010)—a multilocus coalescent-based method that includes prior information about population size and divergence times and uses reversible-jump MCMC to estimate the posterior distribution for different species delimitation models, in the program BPP version 2.0 (Rannala and Yang 2003; Yang and Rannala 2010). We ran the analysis using conditions in Niemiller et al. (2012). We used the *BEAST species tree as the guide tree in each analysis but only included four species at the tips that included *F. agassizii*, *F. papilliferus*, and the two lineages within *Amblyopsis spelaea*. Each analysis was run for 500,000 generations with a burn-in of 50,000 and run at least twice to confirm consistency between runs.

We also assessed the taxonomic distinctiveness of populations north and south of the Ohio River using the genealogical sorting index (*gsi*; Cummings et al. 2008) whereby a quantitative measure of the degree to which ancestry of delimited species is exclusive is generated for individual genes and for multilocus data combined. The relative degree of exclusive ancestry is

on a scale from 0 to 1, where 1 indicates complete monophyly. Using this statistic, hypothesized species can be tested against a null hypothesis of no divergence. We calculated an ensemble *gsi* (*egsi*) and *gsi* for each locus using the Genealogical Sorting Index web server (<http://www.genealogicalsorting.org>). Consensus gene trees were used as input trees. The null hypothesis that the degree of exclusive ancestry is observed by chance alone (i.e., no divergence) was evaluated by estimating a *P* value using 10,000 permutations.

Results

POPULATION SUBDIVISION IN *AMBLYOPSIS*

Estimation of gene genealogies resulted in nearly identical topologies for most loci (Fig. S1), with two major clades in *Amblyopsis* corresponding to samples north and south of Ohio River, respectively. These clades were reciprocally monophyletic for the mitochondrial *nd2* and nuclear *rho* loci. Although reciprocal monophyly could not be tested owing to low differentiation at the other three loci, the two groups did not share *s7* alleles.

Haplotype networks also revealed division of *Amblyopsis* into two groups corresponding to alleles found north and south of the Ohio River (Fig. 2). No haplotypes were shared between groups for *nd2*, *s7*, and *rho*, whereas little variation was found for *rag1* and *tbr1*. Overall, few alleles were observed for nuclear loci (maximum of four in *rag1*). Populations north and south of the

Table 2. Results of partial Mantel tests to test the partial correlation between genetic distance and potential barriers to dispersal after controlling for geographic distance (r). Significant P values are denoted by an asterisk (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$) and indicate that a barrier restricts gene flow between populations of *A. spelaea*.

Barrier	<i>nd2</i>	<i>s7</i>	<i>rag1</i>	<i>rho</i>	<i>tbr1</i>
Ohio River	0.99***	0.65***	0.10	1.00***	0.21
Hydrological subbasins	-0.67***	-0.40***	0.20	-0.67***	0.18
Hydrological subbasins (Blue-Sinking split)	0.04	0.01	0.08	0.04	0.10
Ecoregions	0.11	0.03	0.06	0.12	0.22

Table 3. Hierarchical analysis of molecular variance for the mtDNA *nd2* locus grouped according to (a) region (north and south of Ohio River), (b) hydrological subbasin, and (c) ecoregion (see Table 1). Significance is based on 10,000 permutations: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Source of variation	df	SS	VC	V%	ϕ -Statistics
(A)					
Among regions	1	558.514	15.438	96.74	$\phi_{CT} = 0.967^{***}$
Among populations within regions	14	23.222	0.364	2.28	$\phi_{SC} = 0.700^{***}$
Within populations	56	8.750	0.156	0.98	$\phi_{ST} = 0.990^{***}$
Total	71	590.486	15.958		
(B)					
Among basins	2	246.782	5.347	45.50	$\phi_{CT} = 0.455^*$
Among populations within basins	13	334.954	6.249	53.17	$\phi_{SC} = 0.976^{***}$
Within populations	56	8.750	0.156	1.33	$\phi_{ST} = 0.987^{***}$
Total	71	590.486	11.752		
(C)					
Among ecoregions	1	298.483	8.111	61.83	$\phi_{CT} = 0.618^*$
Among populations within ecoregions	14	283.253	4.851	36.98	$\phi_{SC} = 0.969^{***}$
Within populations	56	8.750	0.156	1.19	$\phi_{ST} = 0.988^{***}$
Total	71	590.486			

Ohio River were separated by three nucleotide substitutions at the *rho* locus and just one substitution at *s7*. Twenty-seven mutational steps separated these groups of populations at the *nd2* locus. The *nd2* network failed to connect at the 95% confidence interval between haplotypes north and south of the Ohio River, so the connect limit was set at 30 to connect populations for visualization purposes only. An insertion shared by all individuals was found in *rho* that is absent in all other amblyopsids. In addition, individuals south of the Ohio River all shared a mutation that resulted in a premature stop codon in the open reading frame.

BARRIERS TO DISPERSAL

Genetic distances between populations are significantly lower on the same side of the river compared to distances between populations located on opposite sides of the river, even after controlling for geographic distance. Uncorrected pairwise sequence divergence for the mitochondrial *nd2* locus between populations on opposite sides of the Ohio River averaged 0.031 ± 0.005 , whereas sequence divergence averaged 0.0012 ± 0.0004 and 0.0006 ± 0.0003 between populations located on the north and south side

of the Ohio River, respectively. Pairwise sequence divergence was considerably lower for all nuclear loci. The greater divergence between populations on opposite sides of the Ohio River was found for the *rho* locus, averaging 0.0050 ± 0.0003 . The results of partial Mantel tests suggest that the Ohio River is a significant barrier to dispersal and restricts gene flow between populations of *Amblyopsis spelaea* north and south of the river after accounting for geographic distance (Table 2). Hierarchical AMOVA of *nd2* revealed that genetic structure is highly correlated with regions north and south of the Ohio River, as 96.7% of variation was partitioned among regions (Table 3). These analyses strongly indicate that the Ohio River has been a significant barrier for cavefish.

Partial Mantel tests failed to detect significant structure across hydrological boundaries or ecoregions (Table 2), which are both possible alternatives to the Ohio River as barriers to dispersal and gene flow. The partial Mantel test for a partial correlation between genetic distance and hydrological subbasins showed a significant negative relationship. Populations that occur within the Blue-Sinking River subbasin, which spans both sides of the Ohio River, drove this negative relationship. When these populations

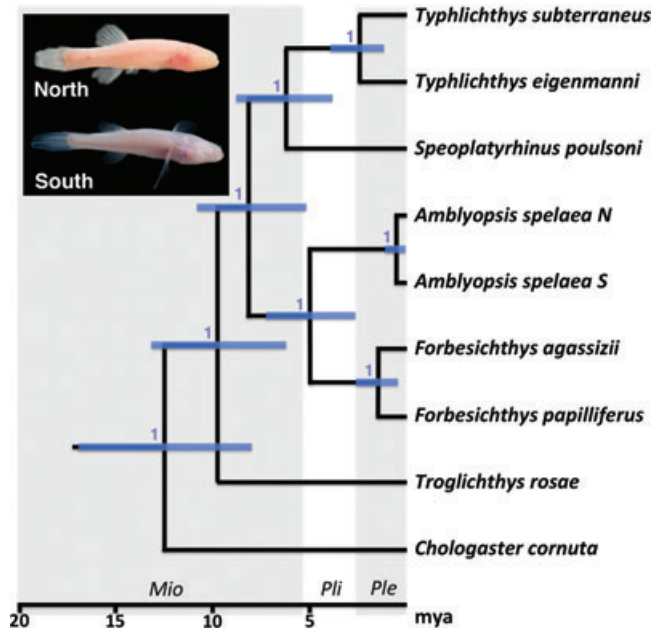


Figure 3. Fossil-calibrated phylogeny for amblyopsid lineages including populations of *Amblyopsis spelaeta* north and south of the Ohio River as separate lineages inferred from the multilocus species tree analysis. Clade posterior probabilities are indicated next to nodes, and uncertainty in divergence time estimates are shown by blue bars on nodes with the length corresponding to the 95% highest posterior density of node ages.

were partitioned into groups north and south of the river, the correlation disappeared and was not significant. Although AMOVAs showed that genetic variation is significantly partitioned among hydrological basins and ecoregions, the percentage of variation explained by partitioning by hydrological basin (45.5%) or ecoregion (61.8%) was substantially less than that when partitioning by region north and south of the Ohio River (Table 3). Given that genetic groups for all three loci with appreciable variation were concordant with the Ohio River break (Fig. 2), no alternative hypotheses are supported by our data.

TIMING OF DIVERGENCE IS CONSISTENT WITH FORMATION OF THE OHIO RIVER

Estimation of the species tree and mean divergence times based on the multilocus dataset revealed a sister relationship between *Amblyopsis* and *Forbesichthys*, which diverged 4.96 mya (95% HPD: 2.64–7.17 mya) during the early Pliocene (Fig. 3), as found previously (Niemiller et al. 2013). Diversification within these genera occurred during the Pleistocene. *F. agassizii* and *F. papilliferus* diverged during the early Pleistocene 1.49 mya (95% HPD: 0.46–2.64 mya), whereas the split between *A. spelaeta* populations north and south of the Ohio River was estimated at 0.53 mya (95% HPD: 0.12–1.06 mya) in the Pleistocene (Fig. 3). The 95% highest posterior density for the divergence date of this split in-

Table 4. Genetic diversity and test statistics of selective neutrality within regions of *Amblyopsis spelaeta* populations north and south of the Ohio River for each locus. Statistics were based on 36 individuals sampled for each group (72 individuals total). K = number of unique haplotypes; S = number of segregating sites; π = nucleotide diversity. Significance for neutrality tests were based on 10,000 permutations. * $P < 0.05$.

Locus	K	S	π	Tajima's D	Fu's Fs	R_2
<i>nd2</i>						
North	7	10	0.0012	-1.52*	1.79	0.07
South	4	4	0.0006	0.88	0.58	0.08
All	11	44	0.0160	2.72	15.56	0.19
<i>s7</i>						
North	2	1	0.0001	1.13	1.36	0.16
South	1	0	0.0000			
All	3	2	0.0007	0.49	0.73	0.14
<i>rag1</i>						
North	3	2	0.0001	-1.50*	-2.59*	0.11
South	2	1	0.0001	0.23	0.23	0.10
All	4	3	0.0001	-1.41*	-3.28*	0.06
<i>rho</i>						
North	1	0	0.0000			
South	1	0	0.0000			
All	2	4	0.0025	3.07	7.97	0.25
<i>tbr1</i>						
North	1	0	0.0000			
South	3	3	0.0007	0.78	1.63	0.08
All	3	3	0.0004	1.16	0.75	0.04

cludes the estimated date of formation of the modern course of the Ohio River ca. 0.8 mya (Gray 1991).

EVIDENCE FOR POPULATION SIZE EXPANSION AND PERSISTENCE IN PERIGLACIAL REFUGIA

Measures of genetic diversity were all higher for the group of populations north of the Ohio River closest to the glacial maximum during the Pleistocene than south of the river, with the exception of the *tbr1* locus (Table 4). Test results for population size changes based on summary statistics in *Amblyopsis* showed a trend of population expansion for both groups; however, only Tajima's D and R_2 for *nd2* and Tajima's D and Fu's Fs for *rag1* were significant in the northern group (Table 4). A signal of population expansion was not detected for the combined dataset leaving out *rag1*.

Bayesian skyride plots did not uncover significant effective population size changes for either group of populations, although there is an increasing trend toward the present day in both groups (Fig. 4). The median TMRCA for the northern group was 54,000 years with a maximum TMRCA of 92,000 years, whereas the median TMRCA for the southern group was 20,000 years with a maximum TMRCA of 42,000 years. The median effective

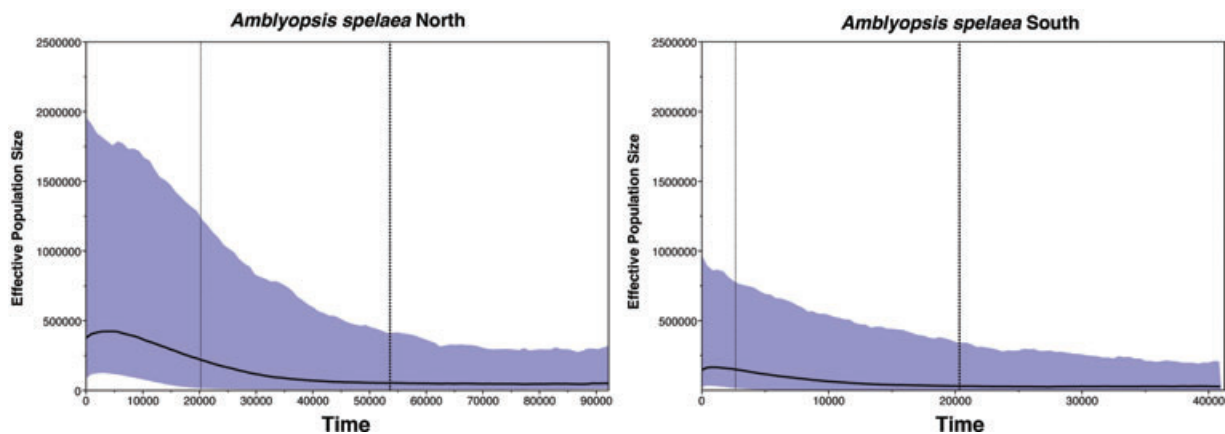


Figure 4. GMRf skyride plots for the *nd2* locus for groups of populations north and south of the Ohio River in *Amblyopsis*. Time (in years) is shown on the x-axis and the effective population size (number of individuals) is shown on the y-axis. The central dark horizontal line in the plot is the median value for effective population size and the shaded area represents the 95% HPD interval for those estimates. The vertical dashed line represents the median TMRCA. The upper 95% HPD for the TMRCA is at the right edge of the plot, whereas the lower 95% HPD is the vertical line to the left of the median.

population size was 380,000 individuals for the northern group but was lower for the southern group at 130,000 individuals. These results indicate that although there is some evidence consistent with population expansion north of the Ohio River, *Amblyopsis* persisted during the most recent glacial cycles in at least two periglacial refugia, including one north of the Ohio River.

Given the low levels of variation in all loci in both northern and southern groups, we have very little power to detect nonequilibrium patterns or demographic differences between groups. However, given the lack of shared alleles at three loci (Fig. 2) and our divergence time estimates (Fig. 3), the alternative hypothesis that northern populations were colonized from the south in the last 20,000 years is not consistent with our data.

CRYPTIC DIVERSITY IN *AMBLYOPSIS*

All species delimitation approaches partitioned populations north and south of the Ohio River into distinct lineages. O'Meara's (2010) nonparametric heuristic approach delimited two species within *Amblyopsis spelaea* with all samples north of the Ohio River in Indiana as a species separate from samples south of the river (Fig. 5). The STRUCTURE analysis also returned $K = 2$ as the optimal value for K , and resulted in complete separation into two groups: a group corresponding to individuals from north of the Ohio River and a group from south of the Ohio River. Estimated proportions of ancestry for all individuals within a group were > 0.98 . Bayesian species delimitation supported the guide tree when assuming four species (two *Amblyopsis* plus *F. agassizii* and *F. papilliferus*), supporting a single model with speciation probabilities of 1.0 for all nodes (Fig. 5). These partitioning methods underscore the pattern illustrated in Fig. 2 and fail to support any additional or alternative patterns of subdivision.

Values of *gsi* and *egsi* indicate a high degree of exclusive ancestry in *A. spelaea* north and south of the Ohio River (Table S2). *Gsi* values were greater than 0.9 for both groups for all loci except *rag1* and *tbr1*, which exhibited shared ancestry of alleles and low levels of variation compared to other loci (Figs. 2, S1). *Egsi* values were also significant for both groups with moderately high values of exclusive ancestry, reflecting lack of differentiation at the *rag1* and *tbr1* loci. These results also are consistent with the Ohio River being a major barrier to dispersal separating *Amblyopsis* into two distinct groups.

Discussion

Determining the factors promoting divergence is often difficult, and is especially problematic for subterranean organisms because of our limited knowledge regarding the delineation of potential vicariant boundaries and connectivity between populations, and the difficulty associated with sampling subterranean habitats. Pleistocene glacial cycles have significantly affected the distribution and diversity of subterranean fauna, as troglobitic species, particularly terrestrial fauna, are largely absent from formerly glaciated regions (Folquier et al. 2008; Culver and Pipan 2010). In this study, we employed multiple, hypothesis-driven approaches to assess the influence of climatic and geological changes during the Pleistocene on a species of amblyopsid cavefish, a top predator in aquatic cave ecosystems in the Interior Plateau of North America. We show that the modern Ohio River has been a significant barrier to dispersal and its formation likely promoted diversification in *Amblyopsis*. We also show that *Amblyopsis* persisted in at least two distinct periglacial refugia, including one north of the Ohio River near the southern glacial maximum during the Pleistocene, rather than isolation in a single, more southern refugium

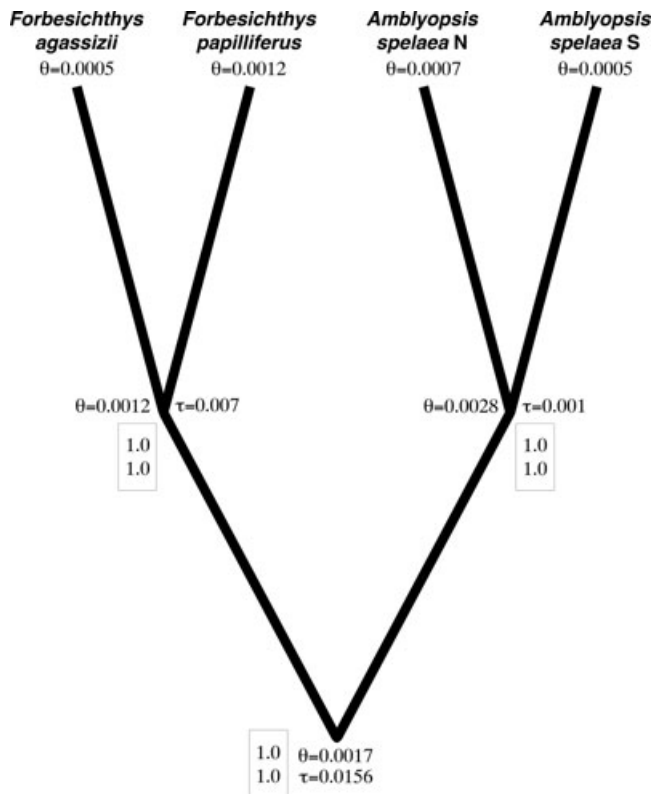


Figure 5. Results from both O’Meara’s (2010) and Bayesian species delimitation support two lineages within *Amblyopsis spelaea*. The percentage of best trees recovering a node in the nonparametric heuristic approach (top) and Bayesian speciation probabilities (bottom) are provided for each node. The Bayesian posterior estimates for θ and τ are also provided on the species tree.

followed by recent northward colonization. Taken together, our results support recognition of two distinct lineages in *Amblyopsis* whose divergence was facilitated by both climatic and geological processes occurring during the Pleistocene.

GEOLOGICAL BARRIERS TO CAVEFISH DISPERSAL

Although large rivers are thought to be significant barriers to dispersal for terrestrial subterranean fauna, it has been hypothesized that major rivers have little influence on the dispersal of many aquatic species, particularly smaller invertebrates (Barr and Holsinger 1985). However, fluvial barriers may be less permeable to larger aquatic fauna, such as crayfishes and cavefishes, which cannot travel through small solution or alluvial channels that likely are filled with sediment and glacial outwash in larger rivers like the Ohio River. Our results for *Amblyopsis* are consistent with this latter hypothesis. Phylogenetic analyses (Figs. 2, S1) strongly suggest that the river is a genetic barrier and limits dispersal between populations north and south of the Ohio River. No alleles were shared on opposite sides of the river for three loci. The presence of shared alleles for *rag1* and *tbr1* likely reflects

retention of ancestral polymorphism, recent divergence, and low mutation rates for these loci.

Divergence dating also supports the Ohio River as a significant isolating barrier, as the 95% highest posterior density for the date of the divergence between populations north and south of the river in *Amblyopsis* (Fig. 3) is consistent with the formation of the modern Ohio River ca. 0.8 mya (Gray 1991). Prior to the river’s formation, the distribution of *Amblyopsis* was likely continuous throughout the cave and karst-bearing geological formations of the Crawford-Mammoth Cave Uplands and Mitchell Plain, as ancient drainages (e.g., the Old Ohio River) were considerably smaller and not as deeply entrenched as the modern Ohio River is today. Once the modern course of the Ohio River was formed after the damming and overflow of headwater drainages of the Teays River (Teller and Goldthwait 1991), it began to cut into the soluble, cave-bearing limestone formations. Entrenchment of the modern Ohio River is attributed to episodic glacial meltwater discharge and probably occurred fairly rapidly during its early development (Teller and Goldthwait 1991) effectively isolating many cave faunas, including *Amblyopsis*.

PERSISTENCE AND DIVERGENCE IN SUBTERRANEAN, PERIGLACIAL REFUGIA

Phylogenetic, divergence dating, and demographic reconstruction analyses support the existence of two refugia during the middle-to-late Pleistocene located on opposite sides of the Ohio River. In addition, our results suggest that subterranean colonization and diversification in *Amblyopsis* was also driven by changing climate during the Pleistocene, consistent with many other subterranean taxa in temperate regions (Barr 1968; Barr and Holsinger 1985; Holsinger 2000; Niemiller et al. 2012). Although determining the exact timing of subterranean colonization is especially difficult for most taxa, several lines of evidence indicate that divergence was facilitated by climatic changes during the Pleistocene. First, the geographic distribution of *Amblyopsis* in close proximity to the glacial maximum implicates a significant influence of Pleistocene climatic fluctuations. It is very unlikely that *Amblyopsis* could have persisted in surface refugia so close to the southern glacial extent due to the harsh periglacial conditions. The onset of glacial advances during the Pleistocene are thought to have occurred rapidly (Adams et al. 1999). Consequently, the ancestor to *A. spelaea* likely was already facultatively living underground but might not have yet become troglomorphic, perhaps much like members of the sister lineage to *Amblyopsis* today (*F. agassizii* and *F. papilliferus*).

Patterns of molecular evolution in the eye photoreceptor gene *rhodopsin* also provide insight into the evolutionary history of *Amblyopsis*. Niemiller et al. (2013) demonstrated that selection is relaxed in subterranean lineages of amblyopsids including *Amblyopsis*. We found that all individuals sampled possess an amino

acid insertion not found in *rhodopsin* gene copies in surface taxa. In addition, all individuals sampled south of the Ohio River possess a mutation that results in a premature stop codon not found in individuals north of the river. The presence of different fixed mutations, including loss-of-function mutations, between populations north and south of the river also suggests that they were isolated in separate refugia throughout the middle to late Pleistocene. The accumulation of nonsynonymous substitutions in *Amblyopsis* is low compared to some other subterranean amblyopsid lineages (Niemiller et al. 2013) and also suggests that subterranean colonization has been relatively recent. Consequently, we cannot rule out the possibility that divergence between populations north and south of the Ohio River occurred before subterranean colonization. Regardless, our study offers strong support for the Ohio River as a barrier facilitating divergence in *Amblyopsis*.

Based on our results, we propose the following hypothesis to explain the evolutionary history of *Amblyopsis* throughout the Pleistocene. Prior to the Pleistocene, the surface ancestor had already migrated into karst regions of the Crawford-Mammoth Cave Uplands and Mitchell Plain inhabiting cool springs, spring runs, and streams, and may also have been present further north than the present-day distribution of *Amblyopsis*. It is also likely that this ancestor was facultatively utilizing caves prior to the Pleistocene. Dramatic climatic shifts and inhospitable surface conditions in the early Pleistocene facilitated further subterranean colonization and extinction of surface populations; however, gene flow between populations continued via subterranean corridors until the overflow of Lake Tight and formation of the modern course of the Ohio River ca. 0.8 mya, which effectively isolated populations to the north and south of this barrier as the river cut through the cave-bearing geological formations. During glacial advances, populations were isolated in refugia and subsequently may have expanded in geographic extent during warmer interglacial periods.

CRYPTIC DIVERSITY AND CONSERVATION IMPLICATIONS

In recent years, phylogeographic studies have uncovered considerable cryptic diversity in several subterranean species groups that exhibit little morphological differentiation and many of which have comparatively large distributions (Finston et al. 2007; Trontelj et al. 2009; Niemiller et al. 2012). We identified two distinct phylogenetic lineages in *Amblyopsis spelaea* that are separated by a major barrier to gene flow and are on independent evolutionary trajectories. These lineages are reciprocally monophyletic at three loci (*nd2*, *s7*, and *rho*) and can be recognized as distinct species under the genealogical species concept (Baum and Shaw 1995) and metapopulation lineage species concept (de Queiroz 1998, 2007). Some morphological data also support recognition of these two lineages in *Amblyopsis* (Poulson 1960), including differences in rudimentary eye size and pigmentation. Based on

these lines of evidence, we believe that populations north of the Ohio River should be recognized as an entity distinct from populations south of the river, which would retain the name *A. spelaea*. However, we refrain from formally describing a new species for the northern lineage until a further examination of morphological differentiation is conducted to help elucidate the taxonomic status of these groups.

The identification of cryptic diversity in *Amblyopsis* has immediate conservation implications. *Amblyopsis spelaea* is listed as “Vulnerable” by IUCN, as regional species of concern by the U.S. Fish and Wildlife Service, and as “Endangered” and “Special Concern” in Indiana and Kentucky, respectively (Niemiller and Poulson 2010). Few cave systems are known to support large populations (>50 individuals), and several populations, particularly those south of the Ohio River, have been significantly impacted by habitat alteration and degradation, groundwater pollution, disease, and over-collection (Pearson and Boston 1995; Niemiller and Poulson 2010). The conservation status of *A. spelaea* in both Kentucky and Indiana may need to be elevated in the future, given evidence of low census population sizes, low genetic diversity, and anthropogenic threats.

Conclusions

Our study highlights the use of multiple, hypothesis-driven approaches in a multilocus framework to elucidate the climatic and geological processes that have influenced the evolutionary history of the only cave-dwelling vertebrate whose distribution extends to the North American glacial maximum of the Pleistocene. Climatic changes have promoted subterranean colonization and speciation, whereas geological processes (i.e., formation of the modern Ohio River) have facilitated isolation and divergence in *Amblyopsis*. Both processes have significantly influenced the levels and structuring of genetic variation. The end result is two phylogenetically distinct lineages that were previously unknown but were diagnosed using multiple species delimitation approaches.

ACKNOWLEDGMENTS

We thank C. Bossu, D. Camacho, J. Grizzle, H. Hensley, P. Hollingsworth, Z. Marion, B. Miller, T. Niemiller, L. Rafai, A. Smith, D. Soares, and T. Webb for assistance with fieldwork. Specimen and tissue collections were authorized by the Indiana Department of Natural Resources and Kentucky Department of Fish and Wildlife Resources (KDFWR). This study was funded by the American Society of Ichthyologists and Herpetologists (M. L. Niemiller), Cave Research Foundation (M. L. Niemiller), KDFWR (contract no. PON2 660 100003354; M. L. Niemiller and B. M. Fitzpatrick), National Science Foundation (DDIG award no. DEB-1011216 to M. L. Niemiller), National Speleological Society (M. L. Niemiller), University of Louisville Multidisciplinary Research Program (J. R. McCandless, C. R. Tillquist, and W. D. Pearson), Department of Ecology and Evolutionary Biology at the University of Tennessee (M. L. Niemiller and B. M. Fitzpatrick), and Yale Institute for Biospheric Studies at Yale University (M. L. Niemiller).

LITERATURE CITED

- Adams, J., M. Maslin, and E. Thomas. 1999. Sudden climatic transitions during the Quaternary. *Prog. Phy. Geogr.* 23:1–36.
- Barr, T. C. 1960. Synopsis of the cave beetles of the genus *Pseudanophthalmus* of the Mitchell Plain in southern Indiana (Coleoptera, Carabidae). *Am. Midl. Nat.* 63:307–320.
- . 1968. Cave ecology and the evolution of troglobites. *Evol. Biol.* 2:35–102.
- Barr, T. C., and J. R. Holsinger. 1985. Speciation in cave faunas. *Annu. Rev. Ecol. Syst.* 16:313–337.
- Baum, D. A., and K. L. Shaw. 1995. Genealogical perspectives on the species problem. Pp. 289–303 in P. C. Hoch and A. G. Stephenson, eds. *Experimental and molecular approaches to plant biosystematics*, Missouri Botanical Garden, St. Louis, Missouri.
- 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. *J. Biogeogr.* 30:1139–1152.
- Bonnet, E., and Y. Van de Peer. 2002. ZT: a software tool for simple and partial Mantel tests. *J. Stat. Soft.* 7:1–12.
- Carstens, B. C., and L. L. Knowles. 2007. Shifting distributions and speciation: species divergence during rapid climate change. *Mol. Ecol.* 16:619–627.
- Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657–1659.
- Cope, E. D. 1872. *Bulletin of the US geological and geographical survey of the territories*, Government Printing Office, Washington, DC, p. 641.
- Culver, D. C., and T. Pipan. 2009. *The biology of caves and other subterranean habitats*. Oxford Univ. Press, Oxford, U. K.
- . 2010. Climate, abiotic factors, and the evolution of subterranean life. *Acta Carsologica* 39:577–586.
- Cummings, M. P., M. C. Neel, and K. L. Shaw. 2008. A genealogical approach to quantifying lineage divergence. *Evolution* 62:2411–2422.
- de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. Pp. 57–75 in D. J. Howard and S. H. Berlocher, eds. *Endless forms: species and speciation*, Oxford Univ. Press, Oxford, U. K.
- . 2007. Species concepts and species delimitation. *Syst. Biol.* 56:879–886.
- Drummond, A., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Eckert, A. J., J. D. Liechty, B. R. Tearse, B. Pande, and D. B. Neale. 2010. DnaSAM: software to perform neutrality testing for large datasets with complex null models. *Mol. Ecol. Res.* 10:542–545.
- Ence, D. D., and B. C. Carstens. 2011. SpedeSTEM: a rapid and accurate method for species delimitation. *Mol. Ecol. Res.* 11:473–480.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14:2611–2620.
- Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* 10:564–567.
- Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.
- Eytan, R. I., and M. E. Hellberg. 2010. Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution* 64:3380–3397.
- Faillie, A., Rivera, I., Deharveng, L., Bourdeau, C., Garnery, L., Queinnec, E., and T. Deuve. 2010. A molecular phylogeny shows the single origin of the Pyrenean subterranean Trechini ground beetles (Coleoptera: Carabidae). *Mol. Phylogenet. Evol.* 54:97–106.
- Finston, T., M. Johnson, W. Humphreys, S. M. Eberhard, and S. A. Halse. 2007. Cryptic speciation in two widespread subterranean amphipod genera reflects historical drainage patterns in an ancient landscape. *Mol. Ecol.* 16:355–365.
- Folquier, A., F. Malard, T. Lefebvre, C. J. Douady, and J. Gibert. 2008. The imprint of Quaternary glaciers on the present-day distribution of the obligate groundwater amphipod *Niphargus virei* (Niphargidae). *J. Biogeogr.* 35:552–564.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Gray, H. H. 1991. Origin and history of the Teays drainage system; The view from midstream. Pp. 43–50 in W. N. Melhorn and J. P. Kempton, eds. *Geology and hydrogeology of the teays-mahomet bedrock valley system*, Special Paper 258, Geological Society of America, Boulder, Colorado.
- Heled, J., and A. Drummond. 2010. Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27:570–580.
- Hewitt, G. M. 2000. The genetic legacy of the quaternary ice ages. *Nature* 405:907–913.
- . 2004. Genetic consequences of climatic oscillations in the Quaternary. *Phil. Trans. Roy. Soc. Lond. B* 359:183–195.
- Hobbs, H. H., H. H. Hobbs III, and M. A. Daniel. 1977. A review of the troglobitic decapod crustaceans of the Americas. *Smithson. Contrib. Zool.* 244:1–183.
- Holsinger, J. R. 1980. *Stygobromus canadensis*, a new subterranean amphipod crustacean (Crangonyctidae) from Canada, with remarks on Wisconsin refugia. *Can. J. Zool.* 58:290–297.
- . 2000. Ecological derivation, colonization, and speciation. Pp. 399–415 in H. Wilkens, D. Culver, and W. Humphreys, eds. *Subterranean ecosystems*. Elsevier, Amsterdam.
- Johnson, N. K., and C. Cicero. 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution* 58:1122–1130.
- Juan, C., and B. C. Emerson. 2010. Evolution underground: shedding light on the diversification of subterranean insects. *J. Biol.* 9:17.
- Juan, C., M. T. Guzik, D. Jaume, and S. J. B. Cooper. 2010. Evolution in caves: Darwin's 'wrecks of ancient life' in the molecular era. *Mol. Ecol.* 19:3865–3880.
- Knowles, L. L., and B. C. Carstens. 2007. Delimiting species without monophyletic gene trees. *Syst. Biol.* 56:887–895.
- Kornobis, E., S. E. Palsson, B. K. Kristjánsson, and J. Svavarsson. 2010. Molecular evidence of the survival of subterranean amphipods (Arthropoda) during Ice Age underneath glaciers in Iceland. *Mol. Ecol.* 19:2516–2530.
- Kozak, K. H., R. A. Blaine, and A. Larson. 2006. Gene lineages and eastern North American paleodrainage basins: phylogeography and speciation in salamanders of the *Eurycea bislineata* species complex. *Mol. Ecol.* 15:191–207.
- Kristjánsson, B. K., and J. Svavarsson. 2007. Subglacial refugia in Iceland enabled groundwater amphipods to survive glaciations. *Am. Nat.* 170:292–296.
- Kubatko, L., H. L. Gibbs, and E. W. Bloomquist. 2011. Inferring species-level phylogenies and taxonomic distinctiveness using multi-locus data in *Sistrurus* rattlesnakes. *Syst. Biol.* 60:393–409.
- Lemmon, E. M., A. R. Lemmon, and D. C. Cannatella. 2007. Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudacris*). *Evolution* 61:2086–2103.
- Lewis, J. J. 1983. The obligatory subterranean invertebrates of glaciated southern Indiana. *Natl. Speleol. Soc. Bull.* 45:34–40.

- Maddison, W. P., and D. R. Maddison. 2005. MacClade: analysis of phylogeny and character evolution. Version 3.0. Sinauer Associates, Inc, Sunderland, Massachusetts.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209–220.
- Mayden, R. L. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Syst. Zool.* 37:329–355.
- McCandless, J. R. 2005. Field surveys and use of a spatially realistic stochastic patch occupancy model to evaluate the conservation status of the northern cavefish, *Amblyopsis spelaea* DeKay. Ph.D. diss., University of Louisville, Louisville, Kentucky.
- McCormack, J. E., J. Heled, K. S. Delaney, A. T. Peterson, and L. L. Knowles. 2011. Calibrating divergence times on species tree versus gene trees: implications for speciation history of *Aphelocoma* jays. *Evolution* 65:184–202.
- Minin, V., E. Bloomquist, and M. Suchard M. 2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* 25:1459–1471.
- Murray, A. M., and M. V. H. Wilson. 1996. A new Paleocene genus and species of percopsiform (Teleostei: Paracanthopterygii) from the Paskapoo Formation, Smoky Tower, Alberta. *Can. J. Earth Sci.* 33:429–438.
- Niemiller, M. L., B. M. Fitzpatrick, and B. T. Miller. 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Mol. Ecol.* 17:2258–2275.
- Niemiller, M. L., B. M. Fitzpatrick, P. Shah, L. Schmitz, and T. J. Near. 2013. Evidence for repeated loss of selective constraint in rhodopsin of amblyopsid cavefishes (Teleostei: Amblyopsidae). *Evolution* 67:732–748.
- Niemiller, M. L., T. J. Near, and B. M. Fitzpatrick. 2012. Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). *Evolution* 66:846–866.
- Niemiller, M. L., and T. L. Poulson. 2010. Studies of the Amblyopsidae: past, present, and future. Pp. 169–280 in E. Trajano, M. E. Bichuette, and B. G. Kapoor, eds. *The biology of subterranean fishes*. Science Publishers, Enfield, New Hampshire.
- O’Meara, B. C. 2010. New heuristic methods for joint species delimitation and species tree inference. *Syst. Biol.* 59:59–73.
- O’Meara, B. C., C. Ane, M. J. Sanderson, and P. C. Wainwright. 2006. Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60:922–933.
- Pearson, W. D., and C. H. Boston. 1995. Distribution and status of the northern cavefish, *Amblyopsis spelaea*. Final report, nongame and endangered wildlife Program, Indiana Department of Natural Resources, Indianapolis.
- Peck, S. B., and K. Christiansen. 1990. Evolution and zoogeography of the invertebrate cave faunas of the Driftless Area of the Upper Mississippi River Valley of Iowa, Minnesota, Wisconsin, and Illinois, USA. *Can. J. Zool.* 68:73–88.
- Pielou, E. C. 1991. *After the ice age: the return of life to glaciated North America*. Univ. of Chicago Press, Chicago, IL.
- Porter, M. L., Dittmar, K., and M. Perez-Losada. 2007. How long does evolution on the troglomorphic form take? Estimating divergence times in *Astyanax mexicanus*. *Acta Carsologica* 36:173–182.
- Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Poulson, T. L. 1960. Cave adaptation in amblyopsid fishes. Ph.D. diss., Univ. of Michigan, Ann Arbor, MI.
- Pritchard, J. K., M. Stephen, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Ramos-Onsins, S., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19:2092–2100.
- Rannala, B., and Z. Yang. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164:1645–1656.
- Ribera, I., J. Fresneda, R. Bucur, A. Izquierdo, A. P. Vogler, J. M. Salgado, and A. Cieslak. 2010. Ancient origin of a Western Mediterranean radiation of subterranean beetles. *BMC Evol. Biol.* 10:29.
- Ribera, I., and A. P. Vogler. 2004. Speciation of Iberian diving beetles in Pleistocene refugia (Coleoptera, Dytiscidae). *Mol. Ecol.* 13:179–193.
- Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- Rosen, D. E., and C. Patterson. 1969. The structure and relationships of the paracanthopterygian fishes. *Bull. Am. Mus. Nat. Hist.* 141:357–474.
- Rosen, D. E. 1962. Comments on the relationships of the North American cavefishes of the family Amblyopsidae. *Amer. Mus. Nov.* 2109:1–35.
- Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst. Zool.* 35:627–632.
- Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *Am. J. Hum. Genet.* 76:449–462.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.
- 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.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Teller, J. T., and R. P. Goldthwait. 1991. The Old Kentucky River: a major tributary to the Teays River. Pp. 29–41 in W. N. Melhorn and J. P. Kempton, eds. *Geology and hydrogeology of the Teays-Mahomet Bedrock Valley System*, Special Paper 258, Geological Society of America, Boulder, Colorado.
- Trontelj, P., C. J. Douady, C. Fiser, J. Gibert, S. Goricki, T. LeFebure, B. Sket, and V. Zakšek. 2009. A molecular test for cryptic diversity in ground water: how large are the ranges of macro-stygobionts? *Freshwater Biol.* 54:727–744.
- Verovnik, R., B. Sket, S. Prevorcnik, and P. Trontelj. 2003. Random amplified polymorphic DNA diversity among surface and subterranean populations of *Asellus aquaticus* (Crustacea: Isopoda). *Genetica* 119:155–165.
- Yang, Z., and B. Rannala. 2010. Bayesian species delimitation using multilocus sequence data. *Proc. Natl. Acad. Sci. USA* 107:9264–9269.

Associate Editor: L. Kubatko

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Loci used in the current study and selected best-fit molecular evolutionary models for character partitions implemented in Bayesian divergence time analyses.

NA – the gene does not contain the specified partition.

Table S2. The genealogical sorting index (*gsi*) for each group of *Amblyopsis spelaea* and each species of *Forbesichthys* for the five loci examined and for all loci combined. Significance is based on 10,000 permutations: * $p < .05$. ** $p < 0.01$. *** $p < 0.001$.

Figure S1. Gene genealogies estimated from partitioned Bayesian analyses of 72 *Amblyopsis* plus outgroup samples for five genes: *nd2*, *s7*, *rag1*, *rho*, and *tbr1*.