Molecular Ecology (2008) **17**, 2258–2275 doi: 10.1111/j.1365-294X.2008.03750.x

Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies

MATTHEW L. NIEMILLER,* BENJAMIN M. FITZPATRICK† and BRIAN T. MILLER*
*Department of Biology, Middle Tennessee State University, Murfreesboro, TN 37132, USA, †Department of Ecology and Evolutionary Biology, University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996, USA

Abstract

Cave organisms occupy a special place in evolutionary biology because convergent morphologies of many species demonstrate repeatability in evolution even as they obscure phylogenetic relationships. The origin of specialized cave-dwelling species also raises the issue of the relative importance of isolation vs. natural selection in speciation. Two alternative hypotheses describe the origin of subterranean species. The 'climate-relict' model proposes allopatric speciation after populations of cold-adapted species become stranded in caves due to climate change. The 'adaptive-shift' model proposes parapatric speciation driven by divergent selection between subterranean and surface habitats. Our study of the Tennessee cave salamander complex shows that the three nominal forms (*Gyrinophilus palleucus palleucus*, *G. p. necturoides*, and *G. gulolineatus*) arose recently and are genealogically nested within the epigean (surface-dwelling) species, *G. porphyriticus*. Short branch lengths and discordant gene trees were consistent with a complex history involving gene flow between diverging forms. Results of coalescent-based analysis of the distribution of haplotypes among groups reject the allopatric speciation model and support continuous or recurrent genetic exchange during divergence. These results strongly favour the hypothesis that Tennessee cave salamanders originated from spring salamanders via divergence with gene flow.

Keywords: cave invasion, divergence with gene flow, *Gyrinophilus*, speciation, subterranean

Received 6 October 2007; revision accepted 11 January 2008

Introduction

The origin of cave species has received considerable attention from biologists attempting to understand ecology and evolution (Barr 1968; Culver 1982; Barr & Holsinger 1985; Holsinger 2000). As with speciation in epigean (surface-dwelling) organisms, the roles of selection, gene flow, and geographical isolation during subterranean speciation are controversial and difficult to resolve (Coyne & Orr 2004; Bolnick & Fitzpatrick 2007), but the common ecogeographical context of subterranean speciation generates a common framework from which to ask general questions (Holsinger 2000). The origin of a cave-dwelling lineage always involves an ecological shift and a generally consistent set of character changes including reduction of eyes and pigmentation, enhancement of extra-optic sensory systems, reduction in fecundity and metabolism, and increased longevity (Barr 1968; Culver 1982; Poulson 1985; Borowsky & Wilkens 2002). An important question is whether such forms have originated in complete geographical isolation (allopatric speciation) or in the face of ongoing dispersal across the epigean–subterranean interface (divergence with gene flow, Rice & Hostert 1993).

A traditional view is that cave organisms are isolated and relictual; populations become trapped underground and slowly evolve troglomorphic characters via 'regressive evolution', where structures such as eyes degenerate because selection no longer eliminates mutations that would have been deleterious on the surface (Darwin 1859; Eigenmann 1909; Culver 1982; Romero & Green 2005). However, some authors favour the view that degenerative characters are

Correspondence: Matthew L. Niemiller, Fax: 865-974-3067; E-mail: mniemill@utk.edu

adaptive underground and evolution of cave specialists is no different from evolution of other habitat specialists (Poulson 1963; Culver 1982; Jeffery 2005; Protas *et al*. 2007). In the latter view, subterranean forms can arise via ecological speciation, a mode of divergence that can be rapid and may not require geographical isolation (Schluter 2000; Coyne & Orr 2004; Hendry *et al*. 2007). Ecological divergence with gene flow has become increasingly credible in recent years (Barluenga *et al*. 2006; Panova *et al*. 2006; Savolainen *et al*. 2006; Hendry *et al*. 2007; Quesada *et al*. 2007), but its prevalence is the subject of continuing debate (Coyne & Orr 2004; Bolnick & Fitzpatrick 2007; Coyne 2007).

In the biospeleological literature, specific allopatric speciation and divergence-with-gene-flow scenarios have been described, respectively, as the 'climate-relict' and 'adaptive-shift' hypotheses (Howarth 1973; Holsinger 2000; Rivera *et al*. 2002). Under the climate-relict model, epigean ancestors adapted to cool, moist environments in temperate areas retreated into subterranean habitats in response to climatic fluctuations. As surface conditions became inhospitable, extirpation of epigean populations facilitated allopatric speciation of subterranean populations (Holsinger 1988, 2000; Ashmole 1993). Under the adaptive-shift model, pre-adapted epigean ancestors invaded subterranean habitats to exploit new niches with reduced competition and quickly evolved in sympatry or parapatry with related surface populations (Howarth 1973, 1981; Holsinger 2000). This is a typical two-habitat ecological speciation scenario (Schluter 2000; Coyne & Orr 2004).

While a parapatric 'adaptive-shift' requires divergent natural selection to overcome the homogenizing process of gene flow, allopatric 'climate-relict' divergence may involve strictly neutral changes, adaptive changes that may or may not contribute to the evolution of reproductive isolation, or 'regressive' changes where cave-associated morphologies come about due to loss-of-function mutations that are maladaptive on the surface but selectively neutral underground. Thus, the allopatric climate-relict scenario may or may not also involve adaptive shifts. The key distinction is not adaptation; it is absence or presence of gene flow during divergence.

Here, we investigate the divergence of the Tennessee cave salamander (*Gyrinophilus palleucus*) species complex and its presumed sister species (see Brandon 1966), the spring salamander, *G. porphyriticus*. Spring salamanders are typical plethodontids with a biphasic life cycle where fully aquatic larvae transform into terrestrial adults (Fig. 1). They are found not only in association with surface springs, seepages, and coldwater streams (Bruce 1972; Petranka 1998), but also in caves (Cooper & Cooper 1968; Petranka 1998). Larval *G. porphyriticus* have relatively broad, flat snouts and small eyes, suggesting adaptation to subsurface habitats in the interstitial zone of streambeds (Brandon 1966; Birchfield & Bruce 2000; Bruce 2003). Tennessee cave salamanders express only a paedomorphic life cycle, where the aquatic larvae become large and sexually mature without ever metamorphosing into a terrestrial form (Brandon 1966). In addition to this shift to an entirely aquatic life cycle, Tennessee cave salamanders exhibit morphological traits typical of cave salamanders in other groups, such as *Eurycea* and *Proteus*. These traits include markedly reduced eyes and pigmentation, a broad head with a spatulate snout, and well-developed lateral line sensory systems (Fig. 1).

Current taxonomy recognizes two subterranean species (Miller & Niemiller 2008): the Tennessee cave salamander (*G. palleucus*) from the Central Basin, Eastern Highland Rim, and escarpments of the southern Cumberland Plateau; and the Berry Cave salamander (*G. gulolineatus*) from the Valley and Ridge (Fig. 1). Two subspecies of *G. palleucus* are recognized: the pale salamander (Fig. 1A), *G. p. palleucus*, and the Big Mouth Cave salamander (Fig. 1B), *G. p. necturoides*. For the purposes of understanding the origin of cave-dwelling lineages with troglomorphic traits, the taxonomic status of a group or its resemblance to a particular species definition is not important. The existence of distinct clusters of diversity is important and we treat the above groups as potentially independent subterranean lineages.

The ranges of Tennessee cave salamanders and spring salamanders (*G. porphyriticus*; Fig. 1C), are largely parapatric (Fig. 1; Miller & Niemiller 2008). Brandon (1971) proposed that the subterranean complex evolved from an epigean, metamorphosing ancestor like *G. porphyriticus* during the Pleistocene as climatic conditions isolated peripheral populations of spring salamanders and facilitated allopatric speciation under the climate-relict model. Cave systems inhabited by the *G. palleucus* complex are estimated to be of Pliocene to Pleistocene age (Barr 1961). However, the phylogeny of the complex and its timing of divergence from *G. porphyriticus* have never been investigated.

To address the question of whether cave salamanders arose in allopatry vs. by divergence with gene flow from *G. porphyriticus*, we used population genetic analyses to estimate historic levels of gene flow among taxa and evaluate the influence of geographical structure on gene flow within and between taxa. Distinguishing allopatric speciation from divergence with gene flow is a major challenge for speciation biology (Coyne & Orr 2004; Fitzpatrick & Turelli 2006). Heuristic interpretation of gene trees has proven unsatisfactory because genealogical patterns in recently isolated populations may be qualitatively similar to those produced in the presence of gene flow (Irwin 2002; Hey & Machado 2003). Recent developments in coalescent theory offer promise for making strong inferences from a Bayesian perspective (Nielsen & Wakeley 2001; Hey & Nielsen 2004; Hey 2005). We used the isolation with migration (IM) model to estimate the posterior probability distributions of gene flow parameters given patterns in mitochondrial DNA (mtDNA) and nuclear DNA sequences (Hey & Nielsen

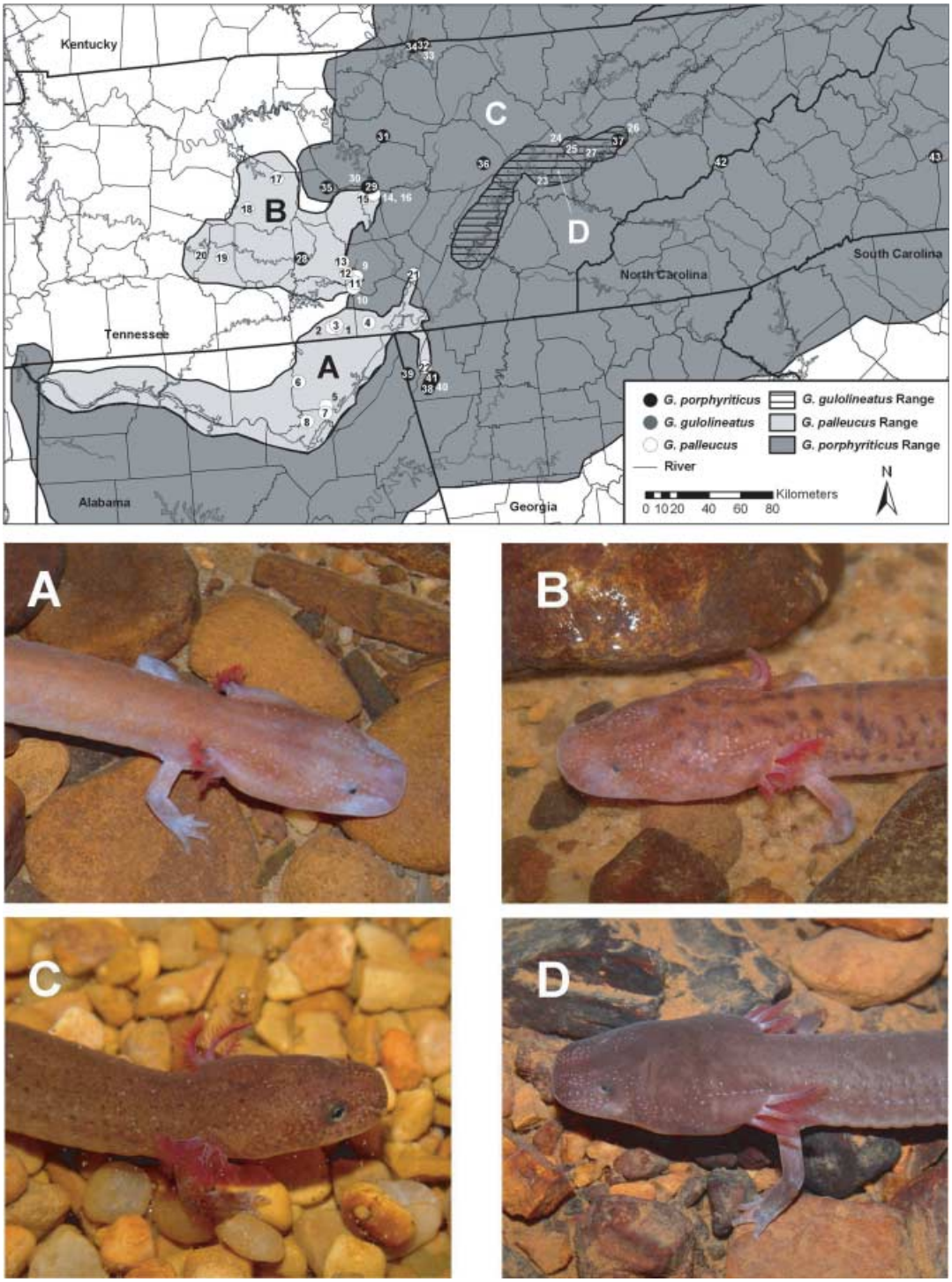

Fig. 1 Sampling localities and distribution (top) of the described taxa within the Tennessee cave salamander complex and the spring salamander (bottom): (A) pale salamander (*G. p. palleucus*), (B) Big Mouth Cave salamander (*G. p. necturoides*), (C) larval spring salamander (*G. porphyriticus*), and (D) Berry Cave salamander (*G. gulolineatus*). Information on numbered localities can be found in Table 1. Note the phenotypic differences between the larval epigean form (C) and the paedomorphic subterranean forms (A, B, and D). Larval subterranean salamanders have reduced eyes and broader, more spatulate snouts than larvae of the epigean form.

2004). Taking the posterior density at or near zero migration as an indication of support for allopatric speciation, our analyses favour the alternative of divergence with gene flow between Tennessee cave salamanders and their epigean sister species, the spring salamander.

Materials and methods

Sampling and molecular methods

As part of a general study of Tennessee cave salamanders, we obtained tail tissue samples from 109 salamanders from 27 localities throughout the range of the Tennessee cave salamander complex and 15 localities of *Gyrinophilus porphyriticus* (Table 1 and Fig. 1; Animal Care and Use Protocol 04-006, Middle Tennessee State University). We were limited to no more than three samples for most localities because of permit restrictions or low abundance. Sampling of *G. porphyriticus* focused on localities in close proximity to Tennessee cave salamander localities because if *G. porphyriticus* is sister to the *G. palleucus* complex, then *G. palleucus* likely diverged from populations of *G. porphyriticus* inhabiting the Cumberland Plateau and Valley and Ridge. Voucher specimens were deposited into the herpetological collection at Middle Tennessee State University.

DNA was extracted from tail tissue and polymerase chain reaction (PCR) was used to amplify portions of two mitochondrial genes, ~850 bp of the 12S ribosomal DNA (rDNA) and 783 bp of cytochrome *b* (cyt *b*), and one nuclear gene, 521 bp of recombination activating gene 1 (RAG-1). This nuclear gene has been utilized with considerable resolving power at many phylogenetic scales in plethodontid salamanders including within genera (Chippindale *et al*. 2004; Min *et al*. 2005; Wiens *et al*. 2006). The 12S rDNA fragment was amplified using primers 12SZ-L and 12SK-H (Goebel *et al*. 1999; Whiting *et al*. 2003) and the cyt *b* fragment was amplified using primers MVZ15 and MVZ16 (Moritz *et al*. 1992) under standard cycling conditions. The RAG-1 fragment was amplified using primers RAGSAL4F (5′-CGTTTCTCYTTCACAYTCATGAC-3′) and RAGSAL3R (5′-GCTGAAAKATCTTYTAYAACTCTG-3′) (P. T. Chippindale, personal communication) using a touchdown protocol. Sequencing reactions were performed using original PCR primers and run on an ABI PRISM 3100 (Applied Biosystems) at GenHunter Corporation.

Forward and reverse sequences for each sample were aligned and edited using SEQMAN (DNASTAR) with ambiguous base calls verified manually by examining the electropherogram for each sequence. These sequences were aligned to each other and to two outgroup sequences for each locus (*Pseudotriton ruber* and *Stereochilus marginatus*, after Chippindale *et al*. 2004; Mueller *et al*. 2004; Macey 2005). Resulting contigs were aligned using CLUSTAL_X (Thompson *et al*. 1997). Direct sequencing of PCR products for RAG-1 revealed a few nucleotide sites at which individuals were heterozygous. Heterozygosity was rare and most haplotypes could be determined unambiguously. Haplotypes for five heterozygous individuals were inferred following Clark (1990).

Phylogenetic analyses

Estimating gene trees. We used COLLAPSE 1.2 (D. Posada) to parse redundant haplotypes for each data set. Gene trees were constructed using maximum parsimony (MP), maximum likelihood (ML), and Bayesian analyses with mtDNA and RAG-1 data sets analysed separately. In addition, unrooted statistical parsimony networks for the mtDNA and RAG-1 data were constructed in TCS 1.18 (Clement *et al*. 2000). We did not concatenate the mtDNA and RAG-1 sequences because this practice can be misleading for three reasons. First, different alleles or loci sampled from the same individual do not have identical histories and it is misleading to assume that they do (Rosenberg & Nordborg 2002; Hey & Machado 2003). Second, there are over three times as many base pairs in our mtDNA data set; therefore, most of the information in a concatenated sequence would reflect mitochondrial history rather than being a fair average of the two histories. Finally, even in situations where the first two concerns do not apply, concatenated sequences can indicate erroneous support for incorrect species trees (Degnan & Rosenberg 2006; Salter Kubatko & Degnan 2007).

MP and ML analyses were conducted in PAUP* 4.0b10 (Swofford 2002) using a heuristic search with tree-bisection–reconnection, ACCTRAN, and 1000 random-taxon-addition replicates. For MP analysis, all sites were treated as unordered, equally weighted characters with gaps treated as missing data. Confidence at each node was assessed using nonparametric bootstrapping (Felsenstein 1985) based on 1000 pseudoreplicates with 100 random-taxon-addition replicates per pseudoreplicate. The optimal model of sequence evolution for each data set (12S, cyt *b*, 12S+cyt *b*, and RAG-1) was determined using Akaike's information criterion (AIC) implemented in MODELTEST 3.7 (Posada & Crandall 1998). ML analyses were performed under the model of evolution selected for each data set. The hypothesis of monophyly of *G. palleucus* and *G. gulolineatus* was evaluated with the Shimodaira–Hasegawa test in PAUP*.

Bayesian posterior probabilities were estimated in MRBAYES 3.1 (Ronquist & Huelsenbeck 2003). Two independent runs using four Markov chains and temperature profiles at the default setting of 0.2 were conducted for 8 million generations, sampling every 100th generation. MODELTEST selected different models of sequence evolution for 12S and first, second, and third position codons of cyt *b*. Therefore, the combined mtDNA data set was partitioned accordingly and unlinked allowing values for transition/transversion

Table 1 Populations of *Gyrinophilus* sampled with locality information, number of individuals (*n*) sampled from each population, watershed, major drainage, mtDNA haplotype with number of individuals sequenced in parentheses, and RAG-1 haplotype with number of individuals sequenced in parentheses used in this study. Discordance between morphological assignment and haplotype are indicated in bold

Population	Species	Locality (cave no.)*	County	State	*n*	Watershed	Major drainage	mtDNA haplotype	RAG-1 haplotype
1	*G. p. palleucus*	Sinking Cove Cave (TFR25)	Franklin	TN	4	Lower Tennessee	Tennessee	palleucus-01 (1), palleucus-02 (3)	palleucus-06 (8)
2	*G. p. palleucus*	Cave Cove Cave (TFR33)	Franklin	TN	3	Lower Tennessee	Tennessee	palleucus-01 (3)	palleucus-06 (6)
3	*G. p. palleucus*	Custard Hollow Cave (TFR7)	Franklin	TN	4	Lower Tennessee	Tennessee	palleucus-02 (4)	**palleucus-01 (2)**, palleucus-06 (6)
4	*G. p. palleucus*	Shakerag Cave (TMN371)	Marion	TN	1	Lower Tennessee	Tennessee	palleucus-16 (1)	palleucus-06 (2)
5	*G. p. palleucus*	Sauta Cave (AJK50)	Jackson	AL	2	Lower Tennessee	Tennessee	palleucus-12 (2)	palleucus-06 (4)
6	*G. p. palleucus*	McFarland Cave (AJK65)	Jackson	AL	1	Lower Tennessee	Tennessee	palleucus-15 (1)	palleucus-04 (2)
7	*G. p. palleucus*	Gross Skeleton Cave (AJK224)	Jackson	AL	1	Lower Tennessee	Tennessee	palleucus-13 (1)	palleucus-07 (2)
8	*G. p. palleucus*	Guffey Cave (AMD317)	Madison	AL	2	Lower Tennessee	Tennessee	palleucus-11 (2)	palleucus-02 (4)
9	*G. p. necturoides*	Big Mouth Cave (TGD2)	Grundy	TN	11	Elk	Tennessee	palleucus-08 (1), **palleucus-17 (8)**, **palleucus-18 (2)**	palleucus-02 (21), palleucus-03 (1)
10	*G. p. necturoides*	Crystal Cave (TGD10)	Grundy	TN	1	Elk	Tennessee	palleucus-08 (1)	palleucus-02 (2)
11	*G. p. necturoides*	Smith Hollow Cave no. 2 (TGD64)	Grundy	TN	3	Elk	Tennessee	palleucus-10 (3)	palleucus-02 (6)
12	*G. p. necturoides*	Blowing Springs Cave (TCF18)	Coffee	TN	3	Elk	Tennessee	palleucus-03, palleucus-08	palleucus-02 (6)
13	*G. p. necturoides*	Lusk Cave (TCF8)	Coffee	TN	3	Elk	Tennessee	palleucus-09 (3)	palleucus-02 (6)
14	*G. p. necturoides*	Jaco Spring Cave (TWR317)	Warren	TN	1	Collins	Cumberland	palleucus-03 (1)	palleucus-02 (2)
15	*G. p. necturoides*	King Cave (TWR295)	Warren	TN	2	Collins	Cumberland	palleucus-03 (2)	palleucus-02 (4)
16	*G. p. necturoides*	Sugarcookie Cave (TWR301)	Warren	TN	1	Collins	Cumberland	palleucus-04 (1)	palleucus-02 (2)
17	*G. palleucus*	Herring Cave (TRU8)	Rutherford	TN	3	Stones	Cumberland	palleucus-14 (3)	palleucus-02 (6)
18	*G. palleucus*	Snail Shell Cave (TRU16)	Rutherford	TN	2	Stones	Cumberland	palleucus-14 (2)	palleucus-02 (4)
19	*G. palleucus*	Gallagher Cave (TMS23)	Marshall	TN	2	Duck	Tennessee	palleucus-05 (1), palleucus-07 (1)	**palleucus-01 (2)**, palleucus-02 (2)
20	*G. palleucus*	Pompie Cave (TMU19)	Maury	TN	3	Duck	Tennessee	palleucus-05 (1), palleucus-06 (2)	palleucus-02 (6)
21	*G. palleucus*	Stone Cave (TSQ7)	Sequatchie	TN	1	Sequatchie	Tennessee	**palleucus-18 (1)**	**palleucus-05 (2)**
22	*G. palleucus*	Fricks Cave (GWK14)	Walker	GA	1	Lower Tennessee	Tennessee	**palleucus-19 (1)**	palleucus-02 (2)
23	*G. gulolineatus*	Berry Cave (TRN3)	Roane	TN	3	Upper Tennessee	Tennessee	**gulolineatus-01 (3)**	gulolineatus-02 (4), **gulolineatus-03 (2)**
24	*G. gulolineatus*	Aycock Spring Cave (TKN172)	Knox	TN	1	Clinch	Tennessee	gulolineatus-02 (1)	gulolineatus-02 (2)
25	*G. gulolineatus*	Christian Cave (TKN49)	Knox	TN	1	Clinch	Tennessee	gulolineatus-03 (1)	gulolineatus-02 (2)
26	*G. gulolineatus*	Meade Quarry Cave (TKN28)	Knox	TN	3	Upper Tennessee	Tennessee	gulolineatus-04 (3)	**gulolineatus-01 (1)**, gulolineatus-02 (5)
27	*G. gulolineatus*	Mudflats Cave (TKN9)	Knox	TN	3	Upper Tennessee	Tennessee	gulolineatus-05 (3)	gulolineatus-02 (6)
28	*G. porphyriticus*	Davidson Branch	Coffee	TN	1	Duck	Tennessee	**porphyriticus-05 (1)**	**porphyriticus-05 (2)**
29	*G. porphyriticus*	Pauley Cave (TDK95)	DeKalb	TN	2	Caney Fork	Cumberland	porphyriticus-08 (1), porphyriticus-09 (1)	porphyriticus-10 (4)

Table 1 *Continued*

Population	Species	Locality (cave no.)*	County	State	*n*	Watershed	Major drainage	mtDNA haplotype	RAG-1 haplotype
30	*G. porphyriticus*	Gar Island Cave (TDK90)	DeKalb	TN	3	Caney Fork	Cumberland	porphyriticus-10 (2), porphyriticus-11 (1)	**porphyriticus-05 (2)**, porphyriticus-10 (4)
31	*G. porphyriticus*	West Cemetery Cave (TPU418)	Putnam	TN	3	Caney Fork	Cumberland	porphyriticus-13 (2), porphyriticus-14 (1)	porphyriticus-08 (1), porphyriticus-10 (5)
32	*G. porphyriticus*	Marcus Cave (TPI76)	Pickett	TN	3	Obey	Cumberland	porphyriticus-01 (2), porphyriticus-17 (1)	porphyriticus-06 (1), porphyriticus-08 (4), porphyriticus-09 (1)
33	*G. porphyriticus*	Mark Us Cave (TPI77)	Pickett	TN	3	Obey	Cumberland	porphyriticus-01 (3)	porphyriticus-01 (1), porphyriticus-07 (1), porphyriticus-08 (4)
34	*G. porphyriticus*	Ringing Rock River Cave (TPI84)	Pickett	TN	1	Obey	Cumberland	porphyriticus-12 (1)	porphyriticus-08 (2)
35	*G. porphyriticus*	Short Mountain†	Cannon	TN	2	Collins/Stones	Cumberland	porphyriticus-15 (2)	**porphyriticus-05 (2)**, porphyriticus-09 (2)
36	*G. porphyriticus*	Spencer Rock Cave (TCD11)	Cumberland	TN	2	Obed	Tennessee	porphyriticus-16 (2)	porphyriticus-01 (1), porphyriticus-04 (1), porphyriticus-11 (2)
37	*G. porphyriticus*	Cruze Cave (TKN24)	Knox	TN	8	Upper Tennessee	Tennessee	porphyriticus-02 (7), porphyriticus-03 (1)	porphyriticus-01 (16)
38	*G. porphyriticus*	Anderson Spring Cave (GWK46)	Walker	GA	5	Lower Tennessee	Tennessee	porphyriticus-06 (5)	porphyriticus-01 (10)
39	*G. porphyriticus*	Hurricane Cave (GDD62)	Dade	GA	2	Lower Tennessee	Tennessee	porphyriticus-07 (2)	porphyriticus-01 (2), porphyriticus-02 (2)
40	*G. porphyriticus*	Pigeon Cave (GWK57)	Walker	GA	4	Lower Tennessee	Tennessee	porphyriticus-06 (4)	porphyriticus-01 (7), **porphyriticus-05 (1)**
41	*G. porphyriticus*	Pocket Branch	Walker	GA	3	Lower Tennessee	Tennessee	porphyriticus-06 (3)	porphyriticus-01 (5), **porphyriticus-05 (1)**
42	*G. porphyriticus*	Cosby Creek, GSMNP	Cocke	TN	1	Pigeon	Tennessee	porphyriticus-18 (1)	porphyriticus-01 (2)
43	*G. porphyriticus*	South Mountains	Burke	NC	1	Catawba	Catawba	porphyriticus-04 (1)	porphyriticus-03 (2)

Locality information and mtDNA haplotype for population 43 correspond to GenBank Accession no. NC_006341 used in Mueller *et al.* (2004). The RAG-1 haplotype from North Carolina is GenBank Accession no. AY691710 used in Chippindale *et al.* (2004).

*Cave number designated by the Tennessee Cave Survey, Alabama Cave Survey, or Georgia Speleological Survey.

†Two individuals were collected < 1 km apart on Short Mountain on each side of the Collins/Stones River drainage divide.

ratio, proportion of invariable sites, and among-site rate heterogeneity to vary across the data sets during analysis. Random trees were used to begin each Markov chain and a molecular clock was not enforced. The first 1.5 million generations were discarded as 'burn-in' to ensure stationarity after examination of the posterior probability. Bayesian analysis on the RAG-1 data set was performed using the same configuration but with nst = 6 and rates = propinv. The first 1.5 million generations were discarded as 'burn-in'. Samples from the stationary distribution of trees were used to generate 50% majority-rule consensus trees for each locus.

Divergence times. To estimate the timescale of diversification, we used two different divergence rates on the mtDNA data set. The general vertebrate molecular clock of 2% uncorrected sequence divergence per million years was employed as the highest rate estimate and a general poikilotherm molecular clock of 0.5% uncorrected sequence divergence per million years was employed as the lowest rate estimate (Avise *et al*. 1998). Lack of *Gyrinophilus* fossils precludes internally calibrated estimates of divergence times. In addition, more advanced approaches to divergence time estimation are not warranted given our result of widespread genealogical discordance. Divergence time estimates are intended only to give a general idea of the timescale of diversification in Tennessee cave salamanders; as with all such molecular clock estimates, they should be interpreted cautiously.

Estimating species trees. To evaluate monophyly of the Tennessee cave salamander complex, we used three methods of estimating species trees from gene trees (Maddison & Knowles 2006; Knowles & Carsten 2007). The methods of minimizing deep coalescents (Maddison 1997) and grouping based on shallowest divergence (Takahata 1989) assume zero gene flow such that the most recent common ancestors of alleles found in different taxa must predate the divergence of those taxa. The method of clustering based on average patristic distances (Knowles & Carstens 2007) only assumes that gene flow has not degraded the expected correlation between average relatedness of individuals between populations and the population divergence time. These methods are implemented and thoroughly explained in the program MESQUITE (Maddison & Maddison 2006).

Allopatric speciation vs. divergence with gene flow

To address the alternative predictions of the allopatric climate-relict model of speciation vs. the adaptive-shift model, which allows divergence with gene flow (Rice & Hostert 1993), we used recently developed methods for estimating levels of gene flow in a general model of population divergence, the isolation with migration (IM) model (Wakeley & Hey 1998; Nielsen & Wakeley 2001; Hey & Nielsen 2004). After separation from an ancestral population, the IM model includes gene flow with rates m_1 and m_2 for gene flow into populations 1 and 2, respectively. The IM model includes complete isolation (allopatry) as a special case where $m_1 = m_2 = 0$. Complete isolation vs. divergence with gene flow can be compared by fitting the IM model to data and estimating the posterior probability that gene flow has been zero since the time of population splitting (Hey & Nielsen 2004; Won & Hey 2005). We used the program IM (Hey & Nielsen 2004) to estimate scaled effective population sizes, migration rates, and divergence times for pairwise comparisons of the four taxa (total of six comparisons) using the full data set including mtDNA and RAG-1. IM is specifically designed to address nonequilibrium scenarios where haplotype sharing and genealogical discordance may result from retention of ancestral polymorphism in recently diverged lineages in addition to potential ongoing gene flow. In cases of very recent divergence, isolation and gene flow may be difficult to distinguish. In such cases, very flat posterior distributions are expected and the proper interpretation would be that the results are inconclusive.

After preliminary runs to determine appropriate priors for subsequent runs and verify convergence of independent runs, a final run was conducted for each pairwise comparison for a minimum of 9 million generations post burn-in (1 million generations) or until minimum effective sample sizes were over 50 (Hey 2005). Each run included four Metropolis-coupled Markov chains, a linear-heating scheme with the first heating parameter set to 0.05, and a maximum of 10 multiple chain-swapping attempts.

To view the history of gene flow in more detail, we recorded the number and time of migration events for each locus over the course of each simulation (Won & Hey 2005). Inspecting the posterior distribution of migration times allows a qualitative evaluation of two alternative gene flow scenarios: secondary contact and hybridization vs. continuous divergence with gene flow. A high concentration of migration events near the present is consistent with secondary contact and renewed gene flow after a period of allopatric divergence, while a broad distribution of migration times when the Markov chain is sampling from its stationary distribution is consistent with continuous or recurrent gene flow since the time of population splitting. As an additional comparison between continuous gene flow vs. secondary contact, we compared the distribution of discordant nodes in each gene tree to null distributions as follows. First, nodes where ranked according to their cladistic distances from the tips following Barraclough & Vogler (2000). The interspecific gene exchange 'events' were mapped onto the gene trees following Slatkin & Maddison (1989) and Maddison & Maddison (2006), and a node was declared discordant if an exchange was mapped to a

branch arising from it. The ranks of discordant and concordant nodes were compared with the Mann–Whitney *U*-test and *P* values checked by randomization. Under a secondary contact scenario, we expect discordant nodes to be more concentrated toward the tips of the gene trees relative to the randomized distributions.

Although the IM model formally assumes that each population is panmictic, some forms of population structure cause little more than a rescaling of the coalescent process for the metapopulation as a whole (Wakeley 2000, 2004; Nordborg 2001; Wakeley & Aliacar 2001; Lessard & Wakeley 2004). Thus, the IM model probably remains a good approximation (with rescaled parameters) for samples from metapopulation structures resembling the island model or stepping-stone model (particularly when sample sizes for each deme are small), but may suffer substantial distortion under strongly hierarchical population structures where coalescence is more likely within certain clusters of demes (Nordborg 2001; Wakeley & Aliacar 2001). While the IM model is probably robust to population structure with regard to distinguishing gene flow and isolation, estimates of effective population sizes and divergence time will be distorted in unpredictable ways (Whitlock & Barton 1997; Wakeley 2000); therefore, we make no attempt to convert the scaled estimates of population size and time to units of real individuals and years, respectively.

Tests of assumptions. The IM model assumes that there has been no recombination or gene conversion during the genealogical history of a single locus and that the variation in the sample is neutral. We tested for recombination in both the mtDNA and RAG-1 data sets using the DSS method implemented in TOPALI (Milne *et al*. 2004) using a window size of 100 and step size of 2. We used the Hudson–Kreitman–Aguade (HKA) test (Hudson *et al*. 1987) to test for deviation from neutrality. This test evaluates the null hypothesis that patterns of polymorphism and divergence in two genes have been shaped by mutation and drift alone. HKA tests were conducted with 10 000 simulations using the computer program HKA (Jody Hey, Rutgers University). Deviations between observed and expected levels of divergence between each pairwise comparison of taxa in IM were summed across the two loci and the probability from the chi-square distribution was calculated.

To test whether genetic population structure is best described as isolation by distance or as hierarchical subdivision, we used distance-based redundancy analysis (dbRDA, Legendre & Anderson 1999; McArdle & Anderson 2001; Geffen *et al*. 2004) to investigate the joint effects of distance and watershed boundaries on genetic structure in the epigean *G. pophyriticus* and subterranean *G. p. necturoides*. This analysis was not relevant to the other two cave taxa because they were confined to single drainages (Table 1). dbRDA has been used as an alternative to partial Mantel tests, which are inadequate because the *P* values are not always indicative of the true type I error rates (Raufaste & Rousset 2001; see also Castellano & Balletto 2002; Rousset 2002). We used the program DISTLM (Anderson 2004) to perform dbRDA using a second-order polynomial function of latitude and longitude as our distance variable set (Borcard *et al*. 1992). First, the relationship between the DNA distance matrices and the distance variable set was analysed alone using dbRDA with *P* values estimated from 9999 permutations of the distance matrix. Then a set of dummy variables indicating the watershed containing each site was analysed as a predictor variable set with the distance variable set fitted as covariates. We used 9999 permutations of the residual distance matrix to estimate *P* values. Two alternative watershed variable sets were analysed (Table 1), one was a 'major drainage' set (Tennessee and Cumberland; the single Catawba sample was not included in the analyses), and the other was a 'watershed' set including finer scale drainage subdivisions (Table 1).

Results

Gene trees

mtDNA. The 12S rDNA and cyt *b* data sets resulted in nearly identical topologies and the combined mtDNA data set of 1641 bp was used to estimate a final gene tree. GenBank Accession numbers for unique haplotypes can be found online (Table S1, Supplementary material). Absence of premature stop codons in cyt *b*, strong bias against guanine on the light strand, and high transition to transversion substitution ratios indicate that amplified sequences were of mitochondrial origin, not nuclear pseudogenes (Zhang & Hewitt 1996). Maximum-likelihood analysis resulted in a single tree of –ln L 5486.55. Details of the ML model of evolution are available online (Table S2, Supplementary material). Bayesian analysis produced a posterior distribution with a mean –ln L of 5639.69 (SD = 0.19). MP, ML, and Bayesian analyses resulted in nearly identical topologies. All mtDNA haplotypes from the Tennessee cave salamander complex were nested within the gene tree of *G. porphyriticus* haplotypes (Fig. 2). No mtDNA haplotypes were shared between named taxa, but genealogical discordance was common and divergence among haplotypes was shallow (Fig. 2). The ML tree, given the backbone constraint of monophyly of alleles within *Gyrinophilus palleucus* and *G. gulolineatus*, was rejected by the Shimodaira–Hasegawa test (log-likelihood ratio = 238.3, $P < 0.001$).

RAG-1. ML analysis of the 521 bp data set resulted in a single tree of –ln L 973.75. Bayesian analysis produced a posterior distribution with a mean –ln L of 1011.20 (SD = 0.17). MP, ML, and Bayesian analyses resulted in nearly identical

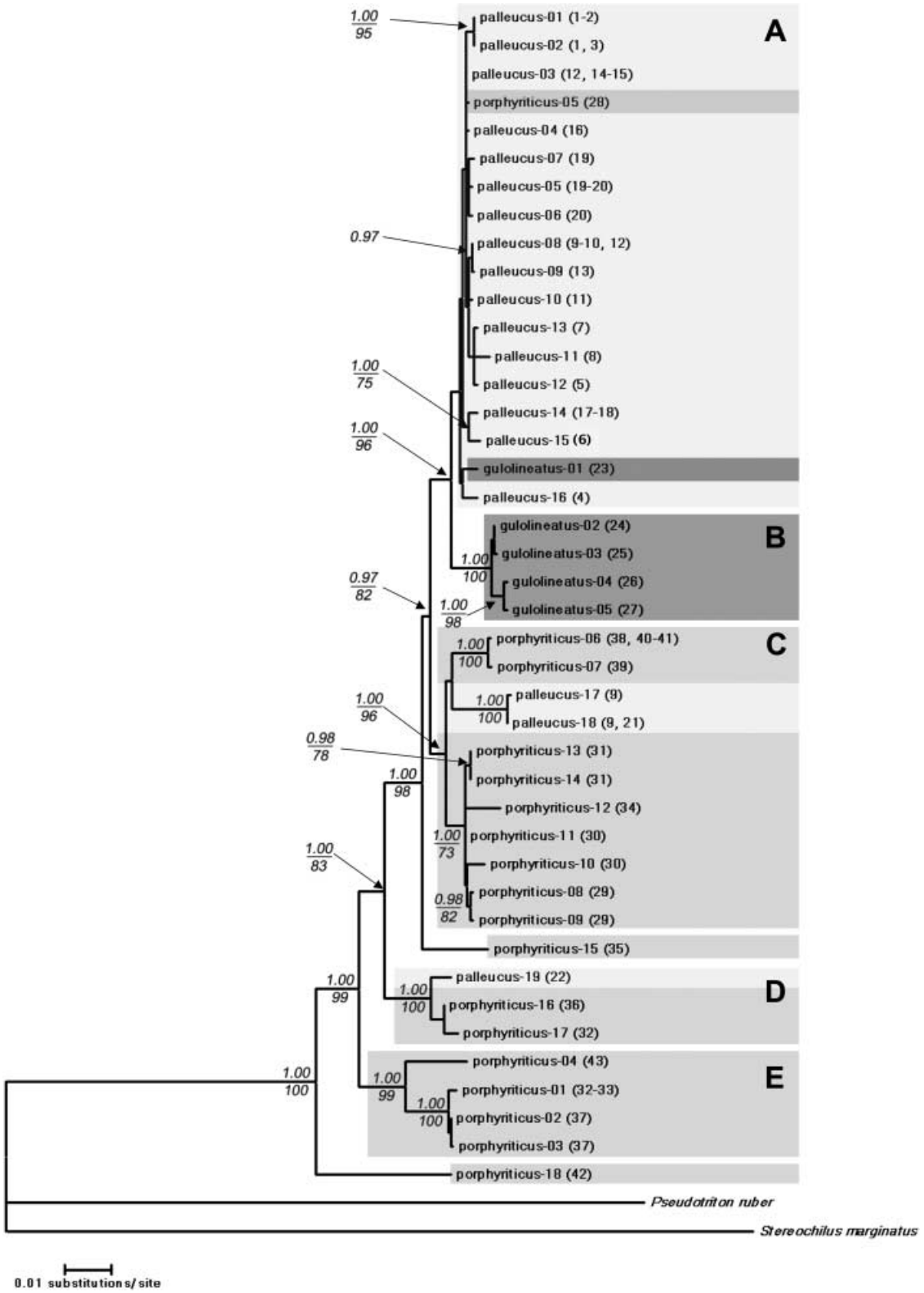

Fig. 2 Maximum-likelihood phylogram of the 1641 bp mtDNA data set (12S rDNA + cyt *b*). Numbers above branches indicate the posterior probability of a clade. Numbers below branches represent bootstrap support for clades resolved in maximum parsimony analysis. Shading represents a *priori* taxonomic designations: light grey, *G. palleucus* (*G. p. palleucus* + *G. p. necturoides*); medium grey, *G. porphyriticus*; and dark grey, *G. gulolineatus*. Numbers in parentheses at the tips indicate population localities as in Table 1 in which the haplotype occurred.

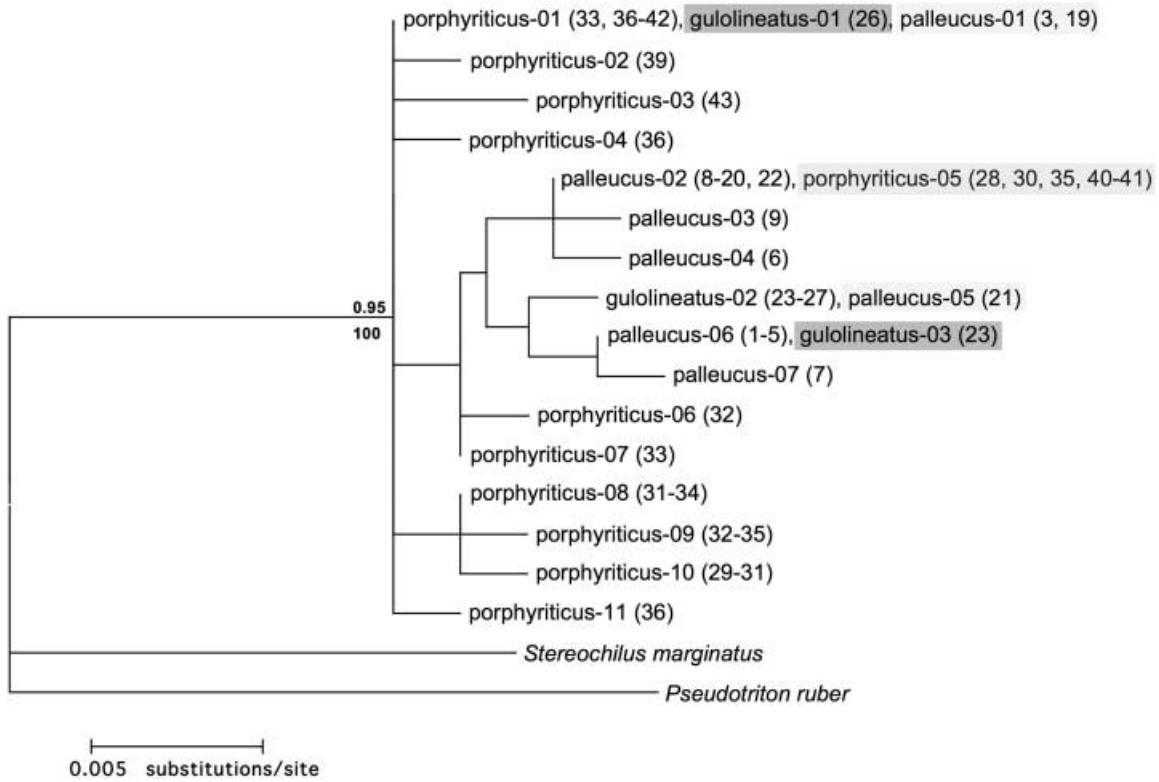

Fig. 3 Maximum-likelihood phylogram of the 521 bp RAG-1 data set. Numbers above branches indicate the posterior probability of a clade. Numbers below branches represent bootstrap support for clades resolved in maximum parsimony analysis. Shading represents discordant haplotypes: light grey, *G. palleucus* (*G. p. palleucus* + *G. p. necturoides*); medium grey, *G. porphyriticus*; and dark grey, *G. gulolineatus*. Numbers in parentheses at the tips indicate population localities as in Table 1 in which the haplotype occurred.

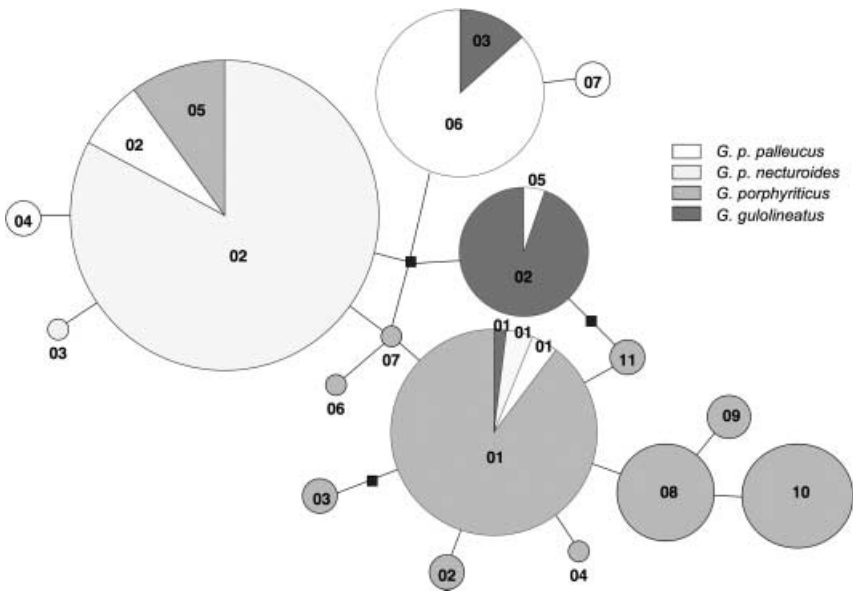

Fig. 4 Unrooted RAG-1 statistical parsimony network. Area of circles is proportional to the number of individuals with that haplotype. Text inside circles corresponds to haplotype number in Table 1 and Fig. 3. Black squares represent unsampled or extinct haplotypes.

topologies. Details of the ML model of evolution are available online (Table S2). The monophyly of *Gyrinophilus* with respect to outgroup haplotypes (*Pseudotriton ruber* and *Stereochilus marginatus*) is supported (PP = 0.95, BP = 100%); however, the order of branching relationships within *Gyrinophilus* could not be elucidated with statistical support because of the limited variation within RAG-1 (Figs 3 and 4). For the most part, populations possessed haplotypes

corresponding to a priori taxonomic designations. Populations of *G. p. palleucus* (populations 1–6) all possessed the same haplotype (palleucus-06) and populations of *G. gulolineatus* (populations 23–27) possessed haplotype gulolineatus-02. Moreover, the majority of *G. p. necturoides* and *necturoides*-like samples including individuals with divergent mtDNA haplotypes possessed haplotype palleucus-02 (populations 9–20). One haplotype was shared among all three species and three other haplotypes were shared between at least two species (Figs 3 and 4). Monophyly of alleles within *G. palleucus* and *G. gulolineatus* was rejected by the Shimodaira–Hasegawa test (log likelihood ratio = 47.3, $P = 0.022$). In addition, the RAG-1 and mtDNA gene trees were significantly different according to the partition-homogeneity test (1000 replicates, $P < 0.001$).

What is the species tree? With the extremely shallow divergence and widespread haplotype sharing evident in our data, a bifurcating species tree could not be resolved unequivocally. Maddison's (1997) approach of choosing the species tree that minimizes the number of deep coalescents in the gene trees supported a sister relationship between *G. palleucus* and the epigean *G. porphyriticus*, implying an earlier, independent origin of the Berry Cave salamander (*G. gulolineatus*). Takahata's (1989) shallowest divergence consistency criterion also favours a slightly more recent split between *G. palleucus* and *G. porphyriticus*. These methods allow for the expectation that recently derived species are genealogically nested within ancestral or sister species gene trees (Takahata 1989; Maddison 1997; Hudson & Coyne 2002). However, they are limited by assuming that genealogical discordance is due solely to incomplete lineage sorting and not gene flow. In contrast, grouping by average sequence divergence, patristic distance, or Φ_{ST} between taxa (Knowles & Carstens 2007) favours monophyly of the Tennessee cave salamander complex: on average, *G. p. palleucus* is most similar to *G. p. necturoides* and *G. gulolineatus* is more similar to *G. palleucus* than to the average *G. porphyriticus*. The latter method fails to appreciate the possibility that the cave-dwelling forms arose independently from the same regional stock, and therefore drew their ancestral haplotypes from a similar nonrandom subset of *G. porphyriticus* gene lineages.

Divergence times. Using sequence divergence between *G. palleucus* mtDNA haplotypes and their nearest *G. porphyriticus* mtDNA haplotypes (after Takahata & Nei 1985), divergence time estimates range from 61 000 to 2.6 million years ago (Pleistocene to mid-Pliocene). This approach gives a range of 244 000 to 2.3 million years ago for the origin of the Berry Cave salamander (*G. gulolineatus*). These ranges are broad, reflecting uncertainty in both the molecular clock and the relationship between gene tree and species tree.

Allopatric speciation vs. divergence with gene flow

The apparent recent origin of the Tennessee cave salamander complex and lack of phylogenetic resolution indicate that a population genetic analysis is the appropriate approach to study the origin of subterranean *Gyrinophilus*. Thus, the remainder of this study focuses on results of fitting the IM model to the data.

Tests of assumptions. Significant recombination was not detected in either mtDNA or RAG-1 data sets using the DSS method as implemented in the program TOPALI. All HKA tests for neutrality for each pairwise comparison were not significant. The test with the largest chi-square value (1.6448) and lowest P value ($P = 0.4394$) was between *G. porphyriticus* and *G. gulolineatus*. Thus, the assumptions of no recombination and no selection were not rejected.

With regard to population structure, all nominal taxa were genetically differentiated (Table 2), supporting treatment of *G. p. necturoides* and *G. p. palleucus* as separate units. Genetic structure was evident within *G. porphyriticus* and *G. p. necturoides* (Table 2). To explicitly test for an effect of watershed or drainage structure over and above the effect of distance, we performed dbRDA on *G. p. necturoides* and *G. porphyriticus*. When spatial variation was taken into account by treating the second-order polynomial function of latitude and longitude as covariables, no association between mtDNA variation and major drainage was detected (Table 3). A weak relationship between genetic variation and finer scale watershed subdivisions was detected for *G. porphyriticus*. Little genetic variation exists for RAG-1 within the named taxa and no association between RAG-1 variation and geographical distance and RAG-1 variation and finer scale watershed subdivisions was detected. However, a significant relationship between RAG-1 variation and major drainage was detected for *G. porphyriticus*.

IM analyses. Given the above results, we considered pairwise comparisons among each of the four described taxa (*G. porphyriticus*, *G. gulolineatus*, *G. p. palleucus*, and *G. p. necturoides*). In so doing, we follow previous authors (e.g. Machado *et al*. 2002; Hey & Nielsen 2004; Hey 2005; Won & Hey 2005) in accepting slight departures from the strict IM model, and our results must be interpreted with these departures in mind. First, the IM model only considers pairs of populations. Analysing multiple pairs is tantamount to assuming that genealogical relationships in each comparison are unaffected by the existence of other populations; high rates of gene flow and dramatic changes in population size would be inconsistent with this assumption. Second, the IM model assumes that each population is panmictic. Significant structure within a population distorts the distribution of coalescent times. As a consequence, estimates

Table 2 Pairwise F_{ST} values between nominate taxa among major drainages. F_{ST} values for mtDNA below diagonal and RAG-1 above

Taxa	*G. porphyriticus* (Tennessee Rr.)	*G. porphyriticus* (Cumberland Rr.)	*G. gulolineatus* (Tennessee Rr.)	*G. p. palleucus* (Tennessee Rr.)	*G. p. necturoides* (Cumberland Rr.)	*G. p. necturoides* (Tennessee Rr.)
G. porphyriticus (Tennessee Rr.)		0.466***	0.786***	0.723***	0.825***	0.824***
G. porphyriticus (Cumberland Rr.)	0.108*		0.687***	0.657***	0.683***	0.744***
G. gulolineatus (Tennessee Rr.)	0.398***	0.392***		0.472***	0.767***	0.799***
G. p. palleucus (Tennessee Rr.)	0.421***	0.416***	0.499***		0.588***	0.651***
G. p. necturoides (Cumberland Rr.)	0.405***	0.400***	0.557***	0.209***		–0.014
G. p. necturoides (Tennessee Rr.)	0.331***	0.312***	0.389**	0.267**	0.274*	

Significance levels: $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

of effective population size and divergence time will also be distorted (Wakeley 2000), and we refrain from presenting or making inferences based on such estimates here. Nevertheless, we assume that the posterior probability of zero gene flow remains a valid test of isolation vs. divergence with gene flow (Machado *et al*. 2002; Hey & Nielsen 2004; Hey 2005; Won & Hey 2005).

Our results clearly reject the simple isolation model ($m_1 = m_2 = 0$). The posterior probability of zero migration between spring salamanders and Tennessee cave salamanders was vanishingly small (Fig. 5A–C). Rather, nonzero gene flow between diverging populations was strongly supported in all comparisons. Estimates of migration rates between taxa indicate asymmetric gene flow. Nonzero gene flow was detected from the epigean *G. porphyriticus* into the subterranean *G. p. palleucus* and *G. gulolineatus* (Fig. 5A, C). However, nonzero gene flow in the opposite direction (from the subterranean to the surface form) was indicated between *G. p. necturoides* and *G. porphyriticus* (Fig. 5B). Posterior distributions indicating gene flow for comparisons among taxa within the Tennessee cave salamander complex have lower resolution as the curves are broad and have higher probabilities at the low limit of resolution (Fig. 5D–F). These results suggest that gene flow between cave-dwelling and surface populations may be greater than gene flow among the three cave-dwelling forms.

Given an inference of nonzero gene flow, an important question is whether that signature results from secondary contact and admixture after a long period of isolation or whether gene flow has been more continuous between the diverging forms. Following Won & Hey (2005), we took an informal approach to this question by recording the timing of simulated migration events in each genealogy sampled by the MCMC algorithm. Reconstructed migration events for comparisons involving *G. porphyriticus* (Fig. 6) for each locus were broadly distributed across time, consistent with continuous or recurrent contact and gene flow between *G. porphyriticus* and the subterranean forms. If nonzero gene flow was due to recent secondary contact, we would expect a greater concentration of migration near the present in Fig. 6. Our analysis of the distribution of discordant nodes also supports continuous gene flow over secondary contact. Mann–Whitney *U*-tests failed to reject random distributions of discordant nodes (for mtDNA W = 81.5, $P = 0.38$; for RAG-1 W = 44.5, $P = 0.99$). In fact, the trend was for gene exchange events (discordant nodes) to be less concentrated toward the tips than expected by chance (observed distributions of discordant nodes were less concentrated toward the tips than 82% of randomized distributions for the mtDNA gene tree and 54% for the RAG-1 gene tree). Secondary contact after allopatric divergence would tend to result in discordant nodes concentrated toward the tips of gene trees.

Table 3 Tests for relationships between genetic variation of *Gyrinophilus* populations and the predictor variables distance, major hydrological drainage, and hydrological watershed, using the dbRDA multivariate *F*-statistic in the DISTLM program. On the left are the results of marginal tests for the distance variable set where a second-order polynomial function of latitude and longitude was fitted. The other two columns show the results of conditional tests evaluating watershed connections as predictors of genetic variation with the distance variable set included as covariables in each analysis. The column '%var' indicates the percentage of the multivariate mtDNA variation explained by the predictor variable. *P* values less than 0.05 are indicated in bold

	Distance			Major drainage			Watershed		
Taxon	F	*P*	%var	F	*P*	%var	F	*P*	%var
mtDNA									
G. p. necturoides	11.835	**0.0001**	89.42	1.738	0.2264	2.38	1.267	0.4058	5.15
G. porphyriticus	5.438	**0.0004**	77.27	0.546	0.6481	1.64	4.174	**0.0340**	19.88
RAG-1									
G. p. necturoides	1.138	0.4511	44.81	0.515	0.5069	4.37	1.386	0.3959	28.13
G. porphyriticus	1.680	0.1723	51.22	9.354	**0.0022**	27.90	2.865	0.0867	40.34

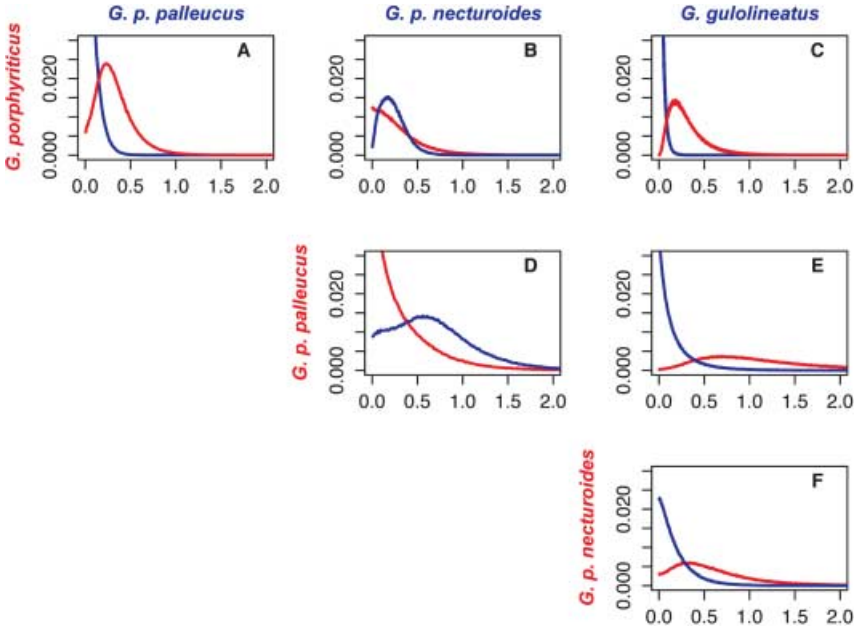

Fig. 5 Posterior distributions of scaled migration rate (m_1 and m_2) estimates for each pairwise comparison among the nominate taxa from the IM program. The *x*-axes are rates of migration per gene copy per mutation event; *y*-axes are marginal posterior probabilities. Posterior distributions shown in red are for migration from the taxon displayed in red font and distributions shown in blue are for migration from the taxon displayed in blue font. The top row shows gene flow between the surface dwelling *G. porphyriticus* and the subterranean *G. p. palleucus* (A), *G. p. necturoides* (B), and *G. gulolineatus* (C). (D), (E), and (F) show gene flow between pairs of subterranean taxa.

Discussion

The transition from epigean to subterranean habitat presents an opportunity for ecological speciation, where divergent selection in alternative habitats leads to evolution of reproductively isolated populations with distinct ecological traits (Schluter 2000). If selection is strong, ecological speciation can proceed without geographical isolation (Rice & Hostert 1993; Schluter 2000; Gavrilets 2003). However, most biospeleologists have followed epigean biologists in favouring allopatric speciation as the primary process in the origin of cave specialists (e.g. Barr & Holsinger 1985). Because subterranean and epigean habitats are often inherently parapatric, allopatric speciation scenarios propose local extinction of epigean populations and isolation of subterranean populations because of climate change creating inhospitable surface conditions. The alternative scenario posits an adaptive shift of subterranean populations in the face of ongoing gene flow with epigean populations (Howarth 1973, 1981). Whether nonallopatric speciation is sympatric or parapatric depends on the initial magnitude of gene flow between habitats (e.g. Gavrilets 2003; Coyne & Orr 2004); either case falls under the more general concept of divergence with gene flow (Rice & Hostert 1993).

Divergence with gene flow has been inferred to explain the origin of some terrestrial troglobites in the tropics (Rivera

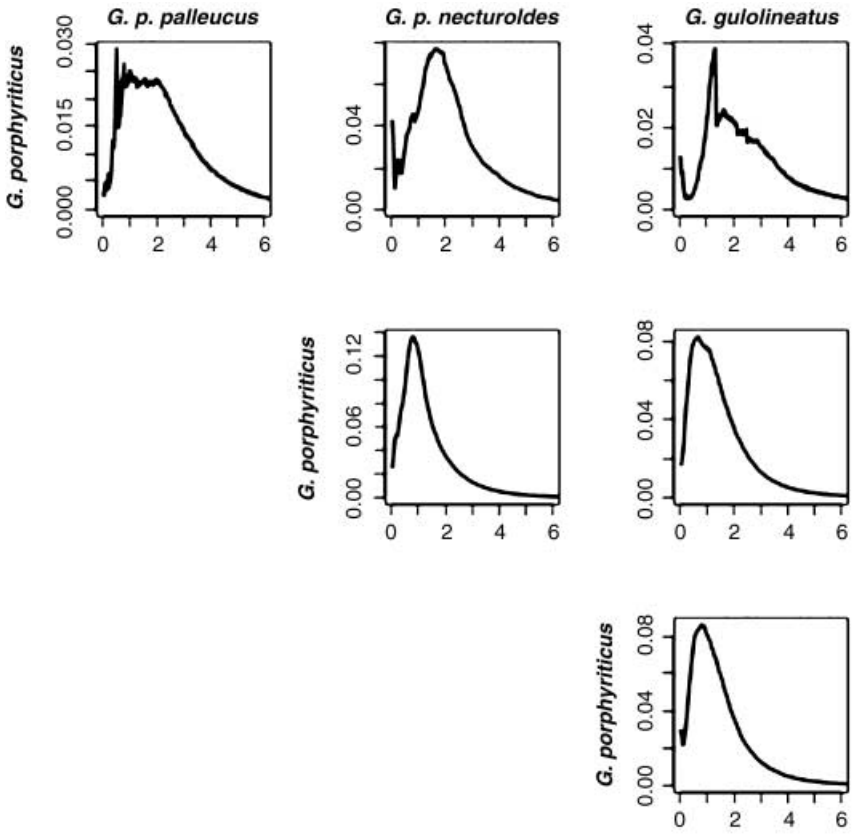

Fig. 6 Distributions of migration events (Won & Hey 2005) for each pairwise comparison among the nominate taxa (*G. p. palleucus*, *G. p. necturoides*, *G. gulolineatus*, and *G. porphyriticus*) from the IM program, summed across loci. The *x*-axes are time (in number of generations ago) multiplied by the effective neutral mutation rate. The *y*-axes are the weighted averages of the individual locus outputs where the weight is by the total number of migration events from the MCMC at stationarity for a locus.

et al. 2002; Schilthuizen *et al*. 2005). Cases from temperate regions involving aquatic organisms are unknown, although the Mexican cavefish, *Astyanax* (Strecker *et al*. 2003), and the temperate amphipod, *Gammarus* (Culver *et al*. 1995), may be cases of divergence with gene flow speciation in progress. Other aquatic subterranean groups in the Appalachian Valley and Interior Plateaus of North America, such as *Orconectes* crayfish (Buhay & Crandall 2005), appear to be more ancient and only distantly related to their nearest epigean ancestors. The ecogeographical context of the origins of such taxa is probably beyond the resolving power of neontological data.

Present-day distributions of *Gyrinophilus porphyriticus* and the Tennessee cave salamander complex are parapatric (Fig. 1). The published distributions of the Berry Cave salamander and *G. porphyriticus* overlap; however, the two species rarely occur in close proximity (Miller & Niemiller 2008). Based on these distributions, Brandon (1971) postulated that the subterranean species evolved from an epigean, metamorphosing ancestor similar to present-day *G. porphyriticus* during the Pleistocene as climatic conditions forced surface populations at the periphery of the species' range underground, thus isolating and facilitating speciation and evolution of troglomorphic characters as predicted by the allopatric, climate-relict model. mtDNA and RAG-1 genealogies support the hypothesis that all three subterranean forms are recently derived from *G. porphyriticus* (Figs 2–4). However, the distribution of polymorphism and divergence in these groups favours divergence with gene flow over allopatric speciation. Posterior distributions of migration rates reject a pure isolation model because the estimated posterior probability of isolation ($m = 0$) between cave and surface lineages is near zero (Fig. 5).

In general, divergence with gene flow can be facilitated by a number of factors including assortative mating, selection against hybrids, or habitat isolation as subterranean populations became more specialized and spread deeper underground (Rivera *et al*. 2002; Coyne & Orr 2004). In *Gyrinophilus*, availability of breeding habitat may be a primary ecological advantage of subterranean colonization. However, permanent residence underground also presents a distinct set of ecological challenges. Survival in caves requires special sensory, metabolic, and life-history adaptations for efficient foraging and resource use (Romero & Green 2005). These adaptations may be particularly likely to entail trade-offs; and trade-offs are necessary for ecological speciation in the face of gene flow (Schluter 2000; Coyne & Orr 2004).

Tennessee cave salamanders show several cave-associated traits not shared by spring salamanders. These include reduced eyes, expanded lateral line systems, and elimination of the terrestrial life-history stage. Reduced eyes may be a correlated response to selection favouring hypertrophy of other sensory systems (Culver *et al*. 1995; Borowsky & Wilkens 2002; Jeffery 2005; Protas *et al*. 2007). Paedomorphosis and a permanently aquatic life cycle in subterranean salamanders is probably adaptive in taking advantage of aquatic resources, while metamorphosis is favoured in the small surface streams inhabited by epigean *Gyrinophilus* (Wilbur & Collins 1973; Whiteman 1994). In other salamanders, paedomorphosis is prevalent in areas where terrestrial habitats are particularly inhospitable (Shaffer & Voss 1996; Chippindale *et al*. 2000; Bonett & Chippindale 2006).

Paedomorphosis may contribute to premating isolation; paedomorphs must court and mate underwater while metamorphosed *Gyrinophilus* may court and mate underwater or on land, although only the latter has been observed (Beachy 1997). Courtship and breeding have never been observed for paedomorphic *Gyrinophilus*. However, instances of successful breeding between paedomorphs and metamorphs exist for other salamander species (Semlitsch & Wilbur 1989; Krenz & Sever 1995; Whiteman *et al*. 1999; Denöel *et al*. 2001). Bonett & Chippindale (2006) suggest that paedomorphosis in plethodontid salamanders may affect mate recognition. The majority of metamorphosing plethodontids exhibit terrestrial courtship where pheromones play an intricate role. Reproductive isolation between paedomorphs and metamorphs may occur readily in plethodontids if paedomorphs do not develop the same pheromones as metamorphs (Bonett & Chippindale 2006).

Temperate cave faunas have often been viewed as isolated and relictual, with troglomorphic traits evolving slowly via 'regressive' evolution (Romero & Green 2005). However, several examples illustrate rapid, adaptive divergence of cave populations from epigean ancestors (Chakraborty & Nei 1974; Wilkens & Hüppop 1986; Culver *et al*. 1995). Here, we present evidence from mitochondrial and nuclear gene genealogies consistent with recent divergence with gene flow of Tennessee cave salamanders from surface-dwelling *G. porphyriticus*. The origin of subterranean species takes place in a simple ecological context where strong natural selection may promote local adaptation and ecological speciation, even in the face of extensive gene flow. With growing acceptance of divergence with gene flow speciation among zoologists (Coyne & Orr 2004), increasing volumes of DNA data, and development of statistical tools like the IM model (Hey & Nielsen 2004), we expect discovery of more examples of divergence with gene flow in cave-adapted animals.

Acknowledgements

We are grateful to N. Mann, G. Moni, A. Moni, C. Kerr, H. Garland, J. Jensen, J. Buhay, A. Wynn, J. Douglas, B. Biddix, B. Walter, B. Glorioso, J. Todd, A. Farone, M. Farone, T. Niemiller, R. Wyckoff, C. Davis, J.A. Miller, J.H. Miller, E. Gray, E. Young, L. Faust, A. Moore, G. Wallace, G. Reynolds, and P. Shah for assistance during field sampling or laboratory work. We thank the Tennessee Cave Survey and the Tennessee Department of Environment and Conservation for providing locality data and the state agencies of Alabama, Georgia, and Tennessee for permits. We thank the Southeastern Cave Conservancy, Inc. and private landowners for allowing access to their property. An earlier draft of this manuscript was reviewed by J. Fordyce, J. Holsinger, T. Poulson, K. Kozak, J. Buhay, and G. Reynolds. This work was funded by the Tennessee Wildlife Resources Agency (contract nos ED-04-01467-00 and ED-06-02149-00), the Department of Biology at Middle Tennessee State University, and the Department of Ecology and Evolutionary Biology at the University of Tennessee.

References

Anderson MJ (2004) DISTLM *Version 5: A FORTRAN Computer Program to Calculate a Distance-Based Multivariate Analysis for a Linear Model*. Department of Statistics, University of Auckland, New Zealand.

Ashmole NP (1993) Colonization of the underground environment in volcanic islands. *Mémoires de Biospéologie*, **20**, 1–11.

Avise JC, Walker D, Johns GC (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 1707–1712.

Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, **439**, 719–723.

Barr TC Jr (1961) *Caves of Tennessee*. Bulletin 64. Tennessee Division of Geology, Nashville, Tennessee.

Barr TC Jr (1968) Cave ecology and evolution of troglobites. *Evolutionary Biology*, **2**, 35–102.

Barr TC Jr, Holsinger JR (1985) Speciation in cave faunas. *Annual Review of Ecological Systematics*, **16**, 313–337.

Barraclough TG, Vogler AP (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *American Naturalist*, **155**, 419–434.

Beachy CK (1997) Courtship behavior in the plethodontid salamander *Gyrinophilus porphyriticus*. *Herpetologica*, **54**, 288–296.

Birchfield GL, Bruce RC (2000) Morphometric variation among larvae of four species of lungless salamanders (Caudata: Plethodontidae). *Herpetologica*, **56**, 332–342.

Bolnick DI, Fitzpatrick BM (2007) Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution and Systematics*, **38**, 459–487.

Bonett RM, Chippindale PT (2006) Streambed microstructure predicts evolution of development and life history mode in the plethodontid salamander *Eurycea tynerensis*. *BioMed Central Biology*, **4**, 6.

Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Borowsky R, Wilkens H (2002) Mapping a cave fish genome: polygenic systems and regressive evolution. *Journal of Heredity*, **93**, 19–21.

Brandon RA (1966) Systematics of the salamander genus *Gyrinophilus*. *Illinois Biology Monographs*, **35**, 1–85.

Brandon RA (1971) North American troglobitic salamanders: some aspects of modification in cave habitats, with special reference to *Gyrinophilus palleucus*. *Bulletin of the National Speleological Society*, **33**, 1–21.

Bruce RC (1972) Variation in the life cycle of the salamander *Gyrinophilus porphyriticus*. *Herpetologica*, **28**, 230–245.

Bruce RC (2003) Ecological distribution of the salamanders *Gyrinophilus* and *Pseudotriton* in a southern Appalachian watershed. *Herpetologica*, **59**, 301–310.

Buhay JE, Crandall KA (2005) Subterranean phylogeography of freshwater crayfishes shows extensive gene flow and surprisingly large population sizes. *Molecular Ecology*, **14**, 4259–4273.

Castellano S, Balletto E (2002) Is the partial Mantel test inadequate? *Evolution*, **56**, 1871–1873.

Chakraborty R, Nei M (1974) Dynamics of gene differentiation between incompletely isolated populations of unequal size. *Theoretical Population Biology*, **5**, 460–469.

Chippindale PT, Price AH, Wiens JJ, Hillis DM (2000) Phylogenetic relationships and systematic revision of central Texas hemidactyliine plethodontid salamanders. *Herpetological Monographs*, **14**, 1–80.

Chippindale PT, Bonett RM, Baldwin AS, Wiens JJ (2004) Phylogenetic evidence for a major reversal in life history evolution in plethodontid salamanders. *Evolution*, **58**, 2809–2822.

Clark AG (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. *Molecular Biology and Evolution*, **7**, 111–122.

Clement MD, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cooper JE, Cooper MR (1968) Cave-associated herpetozoa II: salamanders of the genus *Gyrinophilus*. Alabama caves. *Bulletin of the National Speleological Society*, **30**, 19–24.

Coyne JA (2007) Sympatric speciation. *Current Biology*, **17**, R787–R788.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, Massachusetts.

Culver DC (1982) *Cave Life: Evolution and Ecology*. Harvard University Press, Cambridge, Massachusetts.

Culver DC, Kane TC, Fong DW (1995) *Adaptation and Natural Selection in Caves: the Evolution of* Gammarus minus. Harvard University Press, Cambridge, Massachusetts.

Darwin C (1859) *On the Origin of Species*. John Murray, London.

Degnan JH, Rosenberg NA (2006) Discordance of species trees with their most likely gene trees. *Public Library of Science: Genetics*, **2**, 762–768.

Denöel M, Poncin P, Ruwet J (2001) Sexual compatibility between two heterochronic morphs in the alpine newt, *Triturus alpestris*. *Animal Behavior*, **62**, 559–566.

Eigenmann CH (1909) *Cave Vertebrates of America. A Study in Degenerative Evolution*. Carnegie Institution of Washington, Washington.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Fitzpatrick BM, Turelli M (2006) The geography of mammalian speciation: mixed signals from phylogenies and range maps. *Evolution*, **60**, 601–615.

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

Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology*, **13**, 2481–2490.

Goebel AM, Donelly JM, Atz ME (1999) PCR primers and amplification methods for 12S ribosomal DNA, the control region, cytochrome oxidase I, and cytochrome *b* in bufonids and other frogs and an overview of PCR primers which have amplified DNA in amphibians successfully. *Molecular Phylogenetics and Evolution*, **11**, 163–199.

Hendry AP, Nosil P, Rieseberg LH (2007) The speed of ecological speciation. *Functional Ecology*, **21**, 455–464.

Hey J (2005) On the number of New World founders: a population genetic portrait of the peopling of the Americas. *Public Library of Science: Biology*, **3**, e193.

Hey J, Machado CA (2003) The study of structured populations — a difficult and divided science has new hope. *Reviews Genetics*, **4**, 535–543.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Holsinger JR (1988) Troglobites: the evolution of cave-dwelling organisms. *American Scientist*, **76**, 85–105.

Holsinger JR (2000) Ecological derivation, colonization, and speciation. In: *Ecosystems of the World*, Vol. 30 (eds Wilkens H, Culver DC, Humphreys WF), pp. 399–415. Elsevier, Oxford.

Howarth FG (1973) The cavernicolous fauna of Hawaiian lava tubes. I. Introduction. *Pacific Insects*, **15**, 139–151.

Howarth FG (1981) Non-relictual terrestrial troglobites in the tropic Hawaiian caves. *Proceedings of the 8th International Congress of Speleology (Bowling Green, Kentucky)*, **2**, 539–541.

Hudson RD, Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution*, **56**, 1557–1565.

Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, **56**, 2383–2394.

Jeffery WR (2005) Adaptive evolution of eye degeneration in the Mexican blind cavefish. *Journal of Heredity*, **96**, 185–196.

Knowles LL, Carstens BC (2007) Estimating a geographically explicit model of population divergence for statistical phylogeography. *Evolution*, **61**, 477–493.

Krenz JD, Sever DM (1995) Mating and oviposition in paedomorphic *Ambystoma talpoideum* precedes the arrival of terrestrial males. *Herpetologica*, **51**, 387–393.

Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, **69**, 1–24.

Lessard S, Wakeley J (2004) The two-locus ancestral graph in a subdivided population: convergence as the number of demes grows in the island model. *Journal of Mathematical Biology*, **48**, 275–292.

Macey JR (2005) Plethodontid salamander mitochondrial genomics: a parsimony evaluation of character conflict and implications for historical biogeography. *Cladistics*, **21**, 194–202.

Machado C, Kliman RM, Markert JM, Hey J (2002) Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and its close relatives. *Molecular Biology and Evolution*, **19**, 472–488.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology*, **46**, 523–536.

Maddison WP, Knowles LL (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, **55**, 21–30.

Maddison WP, Maddison DR (2006) MESQUITE: *A Modular System for Evolutionary Analysis*, Version 1.12, http://mesquiteproject.org.

McArdle BH, Anderson MJ (2001) Fitting multivariate models to semi-metric distances: a comment on distance-based redundancy analysis. *Ecology*, **82**, 290–297.

Miller BT, Niemiller ML (2008) Distribution and relative abundance of Tennessee cave salamanders (Gyrinophilus pallecus and Gyrinophilus gulolineatus) with an emphasis on Tennessee populations. *Herpetological Conservation and Biology*, **3**, 1–20.

Milne I, Wright F, Rowe G *et al*. (2004) TOPALI: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics*, **20**, 1806–1807.

Min MS, Yang SY, Bonett RM *et al*. (2005) Discovery of the first Asian plethodontid salamander. *Nature*, **435**, 87–90.

Moritz C, Schneider CJ, Wake DB (1992) Evolutionary relationships within the *Ensatina eschscholtzi* complex confirm the ring species interpretation. *Systematic Zoology*, **41**, 273–291.

Mueller RL, Macey JR, Jaekel M, Wake DB, Boore JL (2004) Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders: novel insights from complete mitochondrial genome sequences. *Proceedings of the National Academy of Sciences, USA*, **101**, 13820–13825.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

Nordborg M (2001) Coalescent theory. In: *Handbook of Statistical Genetics* (eds Balding DJ, Cannings C, Bishop M), pp. 179–212. Wiley, Chichester, UK.

Panova M, Hollander J, Johannesson K (2006) Site-specific genetic divergence in parallel hybrid zones suggests nonallopatric origin of reproductive barriers. *Molecular Ecology*, **15**, 4021–4031.

Petranka JW (1998) *Salamanders of the United States and Canada*. Smithsonian Institution Press, Washington.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Poulson TL (1963) Cave adaptation in amblyopsid fishes. *American Midland Naturalist*, **70**, 257–290.

Poulson TL (1985) Evolutionary reduction by neutral mutations: plausibility arguments and data from amblyopsid fishes and linyphiid spiders. *Bulletin of the National Speleological Society*, **47**, 109–117.

Protas M, Conrad M, Gross JB, Tabin C, Borowsky R (2007) Regressive evolution in the Mexican cave tetra, *Astyanax mexicanus*. *Current Biology*, **17**, 452–454.

Quesada H, Posada D, Caballero A, Moran P, Rolan-Alvarez E (2007) Phylogenetic evidence for multiple sympatric ecological diversification in a marine snail. *Evolution*, **61**, 1600–1612.

Raufaste N, Rousset F (2001) Are partial Mantel tests adequate? *Evolution*, **55**, 1703–1705.

Rice WR, Hostert EE (1993) Laboratory experiments on speciation: what have we learned in 40 years? *Evolution*, **47**, 1637–1653.

Rivera MAJ, Howarth FG, Taiti S, Roderick GK (2002) Evolution in Hawaiian cave-adapted isopods (Oniscidea: Philosciidae): vicariant speciation or adaptive shifts? *Molecular Phylogenetics and Evolution*, **25**, 1–9.

Rodriguez F, Oliver JL, Marin A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology*, **142**, 485–501.

Romero A, Green SM (2005) The end of regressive evolution: examining and interpreting the evidence from cave fishes. *Journal of Fish Biology*, **67**, 3–32.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory, and the analysis of genetic polymorphisms. *Nature Reviews Genetics*, **3**, 380–390.

Rousset F (2002) Partial Mantel tests: reply to Castellano and Balletto. *Evolution*, **56**, 1874–1875.

Salter Kubatko L, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, **56**, 17–24.

Savolainen V, Anstett MC, Lexer C *et al*. (2006) Sympatric speciation in palms on an oceanic island. *Nature*, **441**, 210–213.

Schilthuizen M, Cabanban AS, Haase M (2005) Possible speciation with gene flow in tropical cave snails. *Journal of Zoological Systematics and Evolutionary Research*, **43**, 133–138.

Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford, UK.

Semlitsch RD, Wilbur HM (1989) Artificial selection for paedomorphosis in the salamander *Ambystoma talpoideum*. *Evolution*, **43**, 105–112.

Shaffer HB, Voss SR (1996) Phylogenetic and mechanistic analysis of a developmentally integrated character complex: alternate life history modes in ambystomatid salamanders. *American Zoologist*, **36**, 24–35.

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogeny of alleles. *Genetics*, **123**, 603–613.

Strecker U, Bernatchez L, Wilkens H (2003) Genetic divergence between cave and surface populations of *Astyanax* in Mexico (Characidae, Teleostei). *Molecular Ecology*, **12**, 669–710.

Swofford DL (2002) PAUP*: *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Takahata N (1989) Gene genealogy in three related populations — consistency probability between gene and population trees. *Genetics*, **122**, 957–966.

Takahata N, Nei M (1985) Gene genealogy and variance of interpopulation nucleotide differences. *Genetics*, **110**, 325–344.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

Wakeley J (2000) The effects of subdivision on the genetic divergence of populations and species. *Evolution*, **54**, 1092–1101.

Wakeley J (2004) Metapopulations and coalescent theory. In: *Ecology, Genetics, and Evolution of Metapopulations* (eds Hanski I, Gaggiotti O), pp. 175–198. Elsevier, Oxford, UK.

Wakeley J, Aliacar N (2001) Gene genealogies in a metapopulation. *Genetics*, **159**, 893–905.

Wakeley J, Hey J (1998) Testing speciation models with DNA sequence data. In: *Molecular Approaches to Ecology and Evolution*

(eds Schierwater B, DeSalle R), pp. 157–175. Birkhauser-Verlag, Basel, Switzerland.

Whiteman HH (1994) Evolution of facultative paedomorphosis in salamanders. *Quarterly Review of Biology*, **69**, 205–221.

Whiteman HH, Gutrich JJ, Moorman RS (1999) Courtship behavior in a polymorphic population of the tiger salamander, *Ambystoma tigrinum nebulosum*. *Journal of Herpetology*, **33**, 348–351.

Whiting AS, Bauer AM, Sites JW Jr (2003) Phylogenetic relationships and limb loss in sub-Saharan African scincine lizards (Squamata: Scincidae). *Molecular Phylogenetics and Evolution*, **29**, 582–598.

Whitlock MC, Barton NH (1997) The effective size of a subdivided population. *Genetics*, **146**, 723–733.

Wiens JJ, Engstrom TN, Chippindale PT (2006) Rapid diversification, incomplete isolation, and the 'speciation clock' in North American salamanders (genus *Plethodon*): testing the hybrid swarm hypothesis of rapid radiation. *Evolution*, **60**, 2585–2603.

Wilbur HM, Collins JP (1973) Ecological aspects of amphibian metamorphosis. *Science*, **182**, 1305–1314.

Wilkens H, Hüppop K (1986) Sympatric speciation in cave fishes? Studies on mixed population of epi and hypogean *Astyanax* (Characidae, Pisces). *Zeitschrift für Zoologische Systematik und Evolutionsforschung*, **24**, 223–230.

Won YJ, Hey J (2005) Divergence population genetics of chimpanzees. *Molecular Biology and Evolution*, **22**, 297–307.

Zhang D-X, Hewitt GM (1996) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology & Evolution*, **11**, 247–251.

This study represents Matthew Niemiller's master's thesis on the evolutionary history of cave salamanders with Brian Miller. Niemiller's research interests focus on the ecology, population genetics and phylogeography of subterranean fauna with an emphasis on fish and salamanders. Niemiller is currently a doctoral student in the Department of Ecology and Evolutionary Biology at the University of Tennessee. Ben Fitzpatrick is an Assistant Professor in the Department of Ecology and Evolutionary Biology at the University of Tennessee. Fitzpatrick's research interests include the genetics and biogeography of speciation, the importance of local adaptation for both evolution and conservation management, and invasion biology. Brian Miller is a Professor of Biology at Middle Tennessee State University. Miller studies the evolution, ecology and conservation biology of amphibians and reptiles, particularly salamanders.

Supplementary material

The following supplementary material is available for this article:

Table S1 GenBank Accession numbers for unique mtDNA (12S and cyt *b*) and nuclear (RAG-1) haplotypes observed.

Table S2 The optimal model of sequence evolution for each data set (mtDNA and RAG-1) determined in MODELTEST 3.7 and used for ML analyses.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2008.03750.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.