Gene lineages and eastern North American palaeodrainage basins: phylogeography and speciation in salamanders of the *Eurycea bislineata* species complex

KENNETH H. KOZAK, RUSSELL A. BLAINE and ALLAN LARSON

Department of Biology, Washington University, Saint Louis, MO 63130, USA

**Abstract**

Contemporary North American drainage basins are composites of formerly isolated drainages, suggesting that fragmentation and fusion of palaeodrainage systems may have been an important factor generating current patterns of genetic and species diversity in stream-associated organisms. Here, we combine traditional molecular-phylogenetic, multiple-regression, nested clade, and molecular-demographic analyses to investigate the relationship between phylogeographic variation and the hydrogeological history of eastern North American drainage basins in semiaquatic plethodontid salamanders of the *Eurycea bislineata* species complex. Four hundred forty-two sequences representing 1108 aligned bases from the mitochondrial genome are reported for the five formally recognized species of the *E. bislineata* complex and three outgroup taxa. Within the ingroup, 270 haplotypes are recovered from 144 sampling locations. Geographic patterns of mtDNA-haplotype coalescence identify 13 putatively independent population-level lineages, suggesting that the current taxonomy of the group underestimates species-level diversity. Spatial and temporal patterns of phylogeographic divergence are strongly associated with historical rather than modern drainage connections, indicating that shifts in major drainage patterns played a pivotal role in the allopatric fragmentation of populations and build-up of lineage diversity in these stream-associated salamanders. More generally, our molecular genetic results corroborate geological and faunistic evidence suggesting that palaeodrainage connections altered by glacial advances and headwater erosion occurring between the mid-Miocene and Pleistocene epochs explain regional patterns of biodiversity in eastern North American streams.

**Keywords:** Amphibia, Appalachian Mountains, biogeography, drainage history, Plethodontidae, stream capture, vicariance

Received 28 June 2005; revision accepted 24 August 2005

**Introduction**

A continuing challenge in evolutionary biology is to understand population processes that lead to the formation of new evolutionary lineages. When combined with geographic patterns of genetic variation predicted independently from geological (Durand *et al.* 1999; Waters *et al.* 2001; Johnson 2002) and palaeoclimatic sources (Masta 2000; Knowles 2001; Carstens *et al.* 2004), the genealogical history and spatial distributions of gene lineages provide a powerful tool for understanding historical and contemporary factors that influence the genetic compositions of populations, and ultimately, speciation (Hewitt 2001; Templeton 2001).

In eastern North America, populations of stream-dwelling organisms have experienced dramatic geological changes in the major drainage patterns that structure their geographic patterns of gene exchange, setting the stage for divergence and speciation. Drainage patterns west of the Eastern Continental Divide (ECD) (Fig. 1A), a major physiographic feature that separates eastern-flowing drainages of the Atlantic Slope from those that ultimately drain into the Gulf of Mexico, were profoundly different from current ones (Fig. 1B). As a result of Pleistocene glacial advances, flow patterns of rivers were reversed to the south, resulting in fusion of formerly isolated drainage systems and the origin...
of the modern Ohio River basin. Likewise, the contemporary Tennessee River basin is a hypothesized composite of two major drainages basins that had independent southward courses to the Gulf of Mexico prior to the close of the Pliocene epoch.

Similarly, as a result of stream capture, headwaters of Atlantic Slope drainages are composites of formerly separate drainages. Stream capture occurs when differential erosion at the headwaters of a drainage divide transfers part of a stream from one drainage system to another, facilitating biotic exchange between previously isolated drainage basins (Burr & Page 1986). This process permits dispersal between formerly isolated drainages and ultimately may lead to speciation if the new drainage divide continues to impede gene exchange among drainage systems. Both hydrogeology and zoogeographic patterns in freshwater fishes indicate that stream captures diverted the headwaters of the New and Upper Tennessee systems into the James–Roanoke and Santee drainages of the Atlantic Slope, respectively (Ross 1969, 1971; Hocutt et al. 1986; Starnes & Etner 1986).

The complex history of drainage-basin fragmentation and fusion in eastern North America generates a priori hypotheses for the phylogeographic structuring of populations throughout their evolutionary histories. First, the hypothesis that the Upper and Lower portions of the Tennessee River had independent outlets to the Gulf of Mexico predicts that haplotypes sampled from the present-day Tennessee River will not form a clade. Instead, an ancient pre-Pleistocene phylogeographic break is expected among haplotypes sampled from the Upper and Lower Tennessee River basins. Haplotypes sampled from the Lower Tennessee River are expected to group with those from the Cumberland drainage, while those from the Upper Tennessee River are predicted to group with those from Gulf drainage systems. Second, the
hypothesis that stream capture played an important role in
the movement of populations across the ECD predicts that
haplotypes in translocated rivers should retain phylogenetic
affinity with those of their ancestral drainage. Specifically,
apopulations sampled from the James–Roanoke and Santee
drainages on the Atlantic Slope should be derived from
ancestral haplotypes distributed in the New and Upper
Tennessee systems, respectively. As a corollary, given the
hypothesized Pleistocene timing of these stream captures
(Ross 1969, 1971), genetic signatures of range expansion
across the ECD are also predicted. Integration of traditional
molecular-phylogenetic methods and population-genetic
approaches based on coalescent theory permits rigorous
tests of these phylogeographic predictions.

Recent applications of mitochondrial DNA haplotype
phylogenies in North American freshwater fishes demon-
strate the importance of historical drainage connections in
shaping geographic patterns of genetic variation (Strange
2002; Berendzen et al. 2003). However, there have been
surprisingly few tests of drainage evolution as a determinant
of biotic diversity in other codistributed stream-associated
organisms (Routman et al. 1994; Voss et al. 1995). Here, we
take the relationship between the geological history of
eastern North America drainage basins and geographic
patterns of mtDNA sequence variation in semiaquatic pletho-
dontid salamanders of the Eurycea bislineata (two-lined
salamanders) species complex.

Populations of two-lined salamanders are common inhab-
itants of eastern North American streams. Several aspects
of their life history suggest that gene flow occurs primarily
through stream corridors and therefore predict phyloge-
ographic structuring that reflects the hydrogeological history
of drainage basins. Both larvae and adults show seasonal
upstream and downstream movements (Bruce 1986). In some
populations, adults engage in seasonal terrestrial movements;
however, they are highly philopatric and return to their
natal streams to overwinter and to breed (MacCulloch &
Bider 1975). Current taxonomy of the group recognizes five
species distributed across the Interior Lowlands, Appala-
chian Highlands, Piedmont, and Coastal Plain of eastern
North America (Sever et al. 1976; Jacobs 1987; Kozak &
Lanoo 2005), making it an ideal study system for investi-
gating the biological significance of geological interactions
among the region’s major drainage basins. Furthermore,
despite overall morphological stasis, the E. bislineata complex
demonstrates evidence of ancient geographic genetic fragmenta-
tion dating to as early as the Miocene (Jacobs 1987).

If gene exchange among populations is influenced by the
interconnectivity of streams, phylogenetic relationships
among gene lineages and processes generating population
structure within those lineages should reflect historical
connections among drainage systems. In this study, we
use a combination of traditional phylogenetic, matrix-
correspondence, nested clade, and population-demographic
analyses to test a priori genetic predictions of historical
fragmentation and population structure derived from
eastern North America’s dynamic hydrologic history.

Materials and methods

Population sampling, DNA sequencing, and alignment
We obtained mtDNA sequence data from 438 salamanders
sampled from 144 localities across the range of the Eurycea
bislineata complex (Fig. 2). Sample sizes range from 1 to 10
individuals per location (mean = 3). Samples of Eurycea
guttulinea, Eurycea longicauda, and Eurycea quadridigitata
were used as outgroups for phylogenetic analyses. Sample
sizes and sampling locations are listed in Table S1
(Supplementary material).

DNA extraction, amplification, and sequencing were
performed as in Kozak et al. (2005). Amplification and
sequencing of the mitochondrial ND2 gene and the
adjacent tRNA\textsubscript{Trp} gene were conducted using primers
L4437 (5’-AACCTTTCGCCGCCATACC-3’) and H5934
(5’-AGRTGCAATGCTTTTGRTT-3’) and primers
L4882 (5’-TGACAAAAACTAGCC-3’) and H5617a
(5’-AAAATRTCTGRTGTATCTC-3’) were used as
internal sequencing primers (Macey et al. 1997). The protein-
coding ND2 gene was translated to amino acids using
MACCLADE (Maddison & Maddison 1992) to check for
premature stop codons. The trnA sequence was aligned
manually based on models of secondary structure (Macey
& Verma 1997). Alignment was straightforward and un-
ambiguous. Length variation consisted of a single base-pair
insertion/deletion in the D-loop of the trnA\textsubscript{Trp} gene in
some individuals.

Phylogenetic analyses
Hierarchical phylogenetic relationships among haplotypes
were estimated using Bayesian and parsimony optimality
criteria. MR MODELTEST 1.1b was used to select the model of
nucleotide substitution that best fit the mDNA sequence data
(www.ebc.uu.se/systzoo/staff/nylander.html). Bayesian
phylogenetic analyses were then implemented in MRBAYES
3.01 (Huelsenbeck & Ronquist 2001) using a GTR + I + \Gamma
model of evolution. Four incrementally heated Markov
chains were run for 5 × 10\textsuperscript{6} generations, sampling every
5000 generations for a total of 1000 samples. Flat priors
were used for all substitution-parameter estimates, and
random trees were used to begin each Markov chain. To
ensure that the Markov chains reached a stable equili-

rium, ln-likelihood values for sampling points were plotted
against generation time. Equilibrium sample points were
used to generate a 50% majority-rule consensus tree, where
the percentage of samples that recover a particular node

© 2006 Blackwell Publishing Ltd, Molecular Ecology, 15, 191–207
Fig. 2 Geographic distribution of *Eurycea bislineata*-complex sampling locations included in this study. Open (A–E) and shaded (F–M) symbols designate the major clades (northern and southern, respectively) to which haplotypes sampled from that population belong. The mixed circle for location 136 denotes sympatric occurrence of haplotypes from the northern and southern clades. Symbols delimit the geographic distributions of the major haplotype clades denoted in Fig. 4. Lineage A: Potomac and Lower Chesapeake drainages; Lineage B: Potomac, Kanawha, Lower Chesapeake, Monongahela, Lower Hudson, Upper Susquehanna, Oswego, Lower Connecticut, Sac, and Upper Connecticut drainages; Lineage C: Cumberland and Lower Tennessee drainages; Lineage D: Green, Wabash and Lower-Ohio, Kentucky, Licking and Little Miami drainages; Lineage E: New, Greenbriar, Middle-Ohio, Roanoke, James, Santee, Shenandoah, and Upper Tennessee (Holston) drainages; Lineage F: Upper Tennessee (Emory) and South Fork Cumberland drainages; Lineage G: Upper Tennessee (Little Tennessee) drainage; Lineage H: Cahaba drainage; Lineage I: Santee and Roanoke drainages; Lineage J: Upper Tennessee (Nolichucky and French Broad), Santee, Enoree, Cape Fear, and Neuse drainages; Lineage K: Pearl drainage; Lineage L: Pascagoula, Middle-Tennessee-Elk, Apalachicola, Florida-Panhandle Coastal, Ochlockonee, Escambia, Alabama, Altamaha, and Savannah drainages; Lineage M: Upper Tennessee (Hiwassee, Little Tennessee, Tuckasegee), Middle Tennessee-Elk, Coosa-Tallapoosa, Savannah, and Santee drainages.
Drainage-basin history and phylogeographic divergence

The hypothesis that interconnectivity of drainage systems influences patterns of phylogeographic structure predicts that salamander populations residing in different drainage basins should display greater genetic divergence than those distributed within the same one. Given the ancient evolutionary history of the *E. bislineata* complex suggested by allozyme-based genetic distances (Jacobs 1987) and mtDNA sequence divergence, phylogeographic patterns may reflect historical as well as contemporary drainage connections. We used partial matrix correlation tests (PMCT) to evaluate congruence between patterns of phylogenetic divergence and drainage-basin connections (Smouse et al. 1986; Legendre et al. 1994; Thorpe et al. 1996).

PMCTs measure the correlation between a response matrix and multiple independent matrices simultaneously (Fig. 3). Simultaneous consideration of multiple matrices is critical because putative causes of phylogeographic divergence may be either correlated or additive in their effects. For example, as the geographic distance separating populations increases, so does the possibility that the populations being compared are distributed among different drainages (i.e. the geographic distance separating populations and the spatial arrangement of drainage basins are not independent). In addition, under isolation by distance the geographic distance separating populations is a significant predictor of their genetic divergence (Slatkin 1993). Therefore, to test the hypothesis that patterns of phylogenetic divergence are influenced by the interconnectivity of stream drainages, it is necessary to determine whether the degree of genetic divergence among populations residing in different drainage basins is greater than expected from geographic distance alone. Finally, drainage connections from different geological epochs are not independent because while some drainage connections were modified, others remained static, making it necessary to estimate the contributions of historical and contemporary drainage patterns relative to each other.

A phylogenetic-divergence matrix measuring the patristic distances separating haplotypes along the Bayesian consensus phylogram was designated the response matrix; the independent matrices included a matrix of great-circle distances between all pairs of sampling locations to control for the spatial effects mentioned above, binary design matrices corresponding to Pliocene, Pleistocene, and contemporary drainage connections, and the current species-level taxonomy. For the drainage-connection matrices, each element was set to zero if the haplotypes being compared were in the same drainage basin for the time period under consideration, or one if they were in different drainage basins. Similarly, haplotypes sampled from the same or different formally recognized species were assigned a value of zero or one, respectively. Each matrix was normalized (zero mean, unit variance). Because elements within individual matrices are not independent, significance of the partial regression coefficients was assessed by randomly permuting the order of values in the phylogenetic-divergence matrix 9999 times using the software package PERMUTE! (Casgrain 2001). If
fewer than 5% of the randomizations had partial regression coefficients greater than, or equal to those of the real data set, the results were considered significant. Figure 3 illustrates a hypothetical example outlining the basic procedure used to conduct the PMCTs.

**Molecular demographic and nested clade analyses**

A potential limitation of using contemporary distributions of lineages to assess the geographic context of lineage splitting is that ranges may shift over time (Losos & Glor 2003). To assess the geographic stability of lineages, we used Tajima's $D$-test (Tajima 1989) as implemented in Arlequin version 2.0 (Schneider et al. 2000) to test for population-genetic equilibrium within major population-level lineages. Assuming neutrality of molecular variation, this test can be used to detect increases in population size that accompany recent range movements. To test whether patterns of mtDNA variation are consistent with the hypothesis of neutrality, we employed the MK test (McDonald & Kreitman 1991) in Dnasp (Rozas et al. 2003), using haplotypic variation within and among the major lineages resolved in the hierarchical phylogenetic analyses to compare the ratio of nonsynonymous and synonymous substitutions within and among species, respectively.

The hypothesis that dispersal across the ECD was facilitated by headwater stream capture predicts (i) genetic evidence of range expansion across the ECD, and (ii) haplotypes distributed in James–Roanoke and Santee drainages of the Atlantic slope should be derived from those in the New and Upper Tennessee drainages, respectively. These predictions are ideally suited for testing with NCA, which uses the geographic distributions of ancestral (interior) haplotypes relative to younger (tip) ones to draw inferences about processes that have shaped spatial patterns of genetic variation. We used statistical parsimony as implemented in tcs version 1.18 (Clement et al. 2000) to link haplotypes into a minimum-connecting network depicting their genealogical relationships. The procedures of Templeton et al. (1987) and Templeton & Sing (1993) were then used to impose a hierarchical nesting structure on haplotype networks. In cases where it was possible to use a maximum-parsimony tree to link haplotype networks that could not be connected at the 95% level of confidence using statistical parsimony, we grouped individual networks as sister clades at equal nesting levels (Templeton et al. 1987) and...
used outgroup rooting to determine the tip/interior status of the nested groups. Geodis version 2.0 (Posada et al. 2000) was used to calculate (i) the clade distance \( D_c \), which measures the average distance of haplotypes in a group from its geographic centre, (ii) the nested clade distance \( D_n \), which measures how far a haplotype group is from the geographic centre of other groups with which it is nested, and (iii) the average \( D_c \) and \( D_n \), separating interior and tip groupings of haplotypes. To test whether the geographic distributions of haplotypes were more widespread or restricted than expected by chance, we used a categorical permutation contingency analysis. The most likely historical and recurrent processes responsible for statistically significant patterns of phylogeographic variation were inferred using the revised inference key (Templeton 2004).

**Results**

**Hierarchical phylogenetic analyses**

Four hundred forty-two sequences representing 1108 aligned bases are reported for the five formally recognized species of the *Eurycea bislineata* complex and three outgroup taxa. Within the ingroup, 270 haplotypes are recovered from the 144 sampling locations (Table S2). Absence of premature stop codons in the protein-coding ND2 gene region, functional stability of the tRNA\(^{Trp}\) gene, and strong bias against guanine on the light strand indicate that the DNA sequences are from the mitochondrial genome and not nuclear-integrated copies of mitochondrial genes (Zhang & Hewitt 1996). Likelihood-ratio tests favour the GTR + I + \( \Gamma \) model of nucleotide substitution. Including the three outgroup taxa, the haplotype data set contains 632 variable sites of which 509 are parsimony informative (474 within the ingroup). Bayesian phylogenetic analysis using the GTR + I + \( \Gamma \) model of evolution produces a 50% majority-rule consensus tree with a mean ln-likelihood of \(-14382.84\) (SD = 22.93) following a ‘burn in’ of 500 000 generations (Fig. 4). Parsimony analysis results in a single tree of 1301 steps. Because both analyses produced highly congruent estimates of phylogenetic relationship among the major haplotype clades, only the Bayesian consensus phylogram is presented, with posterior probabilities and nonparametric bootstrap values from the parsimony analyses included for shared branches.

*Eurycea bislineata*-complex haplotypes form a well-supported monophyletic group to the exclusion of *Eurycea guttolineata*, *Eurycea longicuda*, and *Eurycea quadridigitata*. Tree-based lineage delimitation identifies 13 phylogenetically and geographically distinct haplotype groups diagnosing putatively independent evolutionary lineages (Fig. 4). In some cases it is possible to delimit additional such haplotype groupings; however, we have conservatively focused on the most inclusive groups of haplotypes that are concordant with geography to delimit population-level lineages. A high ratio of between vs. within-clade molecular divergence indicates a deep history of geographic genetic fragmentation among these lineages: Tamura–Nei distances among lineages average 13.5% (uncorrected average = 11.5%), whereas the mean within-lineage divergence is 1.5% (Table 1). Based on these estimates, the vast majority of divergence events between inferred sister lineages are estimated to have occurred during the late Pliocene to mid-Miocene (\(\sim\)4–11 Myr), with a single lineage-splitting event dating to the Pleistocene (A vs. B: \(\sim\)1.5 Myr). Relative-rate tests suggest that lineage-specific rate heterogeneity is not prevalent in the haplotype data set; only a single lineage (lineage K, \(P < 0.02\)) has accumulated nucleotide substitution at a significantly different rate.

The deepest phylogenetic divergence in the complex separates haplotypes into strongly supported northern and southern clades. Northern clade haplotypes are distributed primarily among drainages in the Lower Tennessee, former Old-Ohio, former Teays, mid-Atlantic slope, and those in the northeast that were inundated by Pleistocene glaciations. Southern clade haplotypes are distributed primarily in the Gulf Coastal Plain, southern Atlantic slope, and Upper Tennessee drainage systems. Haplotypes from the northern and southern clades occur sympatrically in the headwaters of the Santee drainage system in the southern Appalachian Highlands.

In general, major haplotype clades in the *E. bislineata* complex exhibit greater concordance with geography than they do with current taxonomy. Three clades correspond to formally recognized species (A + B: *E. bislineata*; G: *Eurycea junaluska*; H: *Eurycea aquatica*). However, haplotypes sampled from *Eurycea cirrigera* and *Eurycea wilderana* do not form monophyletic groups of haplotypes. Instead, they span the basal north–south split in the phylogeny, interdigitating with each other and the remaining taxa of the complex.

**Drainage-basin history and phylogeographic divergence**

Multiple regression of the five matrices on the haplotype phylogeny was highly significant (Table 2). The partial regression coefficients represent the slope for that variable when all other variables are held constant. Significant partial regression coefficients for geographic distance, Tertiary, Pleistocene, and contemporary drainage connections, and the current alpha taxonomy indicate that each of these factors explains phylogeographic structure in the haplotype phylogeny. Although phylogenetic divergence is strongly correlated with geography, historical drainage-basin connections explain more variation in the structure of the phylogeny than expected from the geographic distance separating haplotype sampling locations. Of the matrices considered, Pliocene drainage-basin connections explain the greatest amount of variation in the phylogeny. Interestingly,
Fig. 4 Bayesian 50% majority-rule consensus phylogram for the 270 *Eurycea bislineata*-complex haplotypes and three outgroup taxa. Posterior probabilities based on 4500 trees (which had a mean ln-likelihood of $-14.382.84$; SD = 22.93) are shown above the branches; nonparametric bootstrap proportions, followed by decay indices for the parsimony analysis are below. Haplotypes in bold are included in the parsimony analyses. Clades for which nested clade analysis (NCA) was performed are shaded grey. For illustrative purposes, branch-support values are omitted from short near-terminal branches within major population lineages. Haplotypes are labelled with either the outgroup species name, or designation listed in Table S2. Locations from which haplotypes were sampled (Table S1; Fig. 2) are given in parentheses. Geographic distributions for major population lineages are depicted using symbols as in Fig. 2. (A) The northern clade containing major population lineages A–E; (B) The southern clade containing major population lineages F–M. Symbols for population lineages A–M follow Fig. 2.

© 2006 Blackwell Publishing Ltd, Molecular Ecology, 15, 191–207
Fig. 4 Continued
current species-level taxonomy. A significant correlation between phylogenetic structure and partial-regression coefficients and associated *Partial regression coefficient; †Student’s $t$-statistic‡ Significance of the partial-regression coefficients and associated $t$-statistics determined by 9999 randomizations of the haplotype patristic-distance matrix.

contemporary drainage connections are negatively correlated with the patristic distance separating lineages, which results from some very divergent haplotype clades occupying the same present-day drainage systems. This phylogenetic pattern is expected when modern drainage basins are composites of previously isolated stream systems. Three formally recognized species in the complex correspond to monophyletic groups of haplotypes (E. bislineata, E. aquatica, and E. jinhaluska), which results in a weak but significant correlation between phylogenetic structure and current species-level taxonomy.

Molecular-demographic and nested clade analyses

Genetic signatures of geographically extensive population expansions are detected in two lineages (B and D) that were closest to the receding Wisconsin Ice Sheet (Table 3). The range of lineage B encompasses northeastern drainage systems that were covered by the Wisconsin Ice Sheet during the last glacial maximum. Similarly, lineage D is distributed across tributaries of the former Old-Ohio River and a single tributary of the ancient Teays River basin (Kentucky River), which were impounded during the last glacial maximum. In contrast, genetic evidence for long-term persistence of lineages and subclades is detected across the Lower Tennessee and Cumberland drainage basins (lineage C), tributaries of the former Teays River basin (lineage E), Gulf Coastal Plain (lineage L), and large portions of the Upper Tennessee River basin (lineage M).

Table 1 Matrix of haplotypic divergences within and between 13 population-level lineages identified through phylogeographic analysis. Average Tamura–Nei-corrected distances between lineages are below the diagonal; average uncorrected proportion of sites differing between lineages are above the diagonal; mean Tamura–Nei distances among haplotypes within inferred population-level lineages are shown in boldface on the diagonal.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.005</td>
<td>0.022</td>
<td>0.108</td>
<td>0.106</td>
<td>0.111</td>
<td>0.127</td>
<td>0.136</td>
<td>0.134</td>
<td>0.136</td>
<td>0.133</td>
<td>0.146</td>
<td>0.129</td>
<td>0.127</td>
</tr>
<tr>
<td>B</td>
<td>0.022</td>
<td>0.002</td>
<td>0.10</td>
<td>0.099</td>
<td>0.105</td>
<td>0.125</td>
<td>0.131</td>
<td>0.129</td>
<td>0.126</td>
<td>0.127</td>
<td>0.138</td>
<td>0.124</td>
<td>0.121</td>
</tr>
<tr>
<td>C</td>
<td>0.111</td>
<td>0.111</td>
<td>0.037</td>
<td>0.128</td>
<td>0.123</td>
<td>0.147</td>
<td>0.149</td>
<td>0.150</td>
<td>0.152</td>
<td>0.154</td>
<td>0.153</td>
<td>0.147</td>
<td>0.144</td>
</tr>
<tr>
<td>D</td>
<td>0.118</td>
<td>0.109</td>
<td>0.146</td>
<td>0.007</td>
<td>0.065</td>
<td>0.139</td>
<td>0.146</td>
<td>0.148</td>
<td>0.147</td>
<td>0.141</td>
<td>0.145</td>
<td>0.142</td>
<td>0.138</td>
</tr>
<tr>
<td>E</td>
<td>0.125</td>
<td>0.117</td>
<td>0.139</td>
<td>0.069</td>
<td>0.028</td>
<td>0.141</td>
<td>0.145</td>
<td>0.140</td>
<td>0.140</td>
<td>0.137</td>
<td>0.143</td>
<td>0.134</td>
<td>0.128</td>
</tr>
<tr>
<td>F</td>
<td>0.145</td>
<td>0.143</td>
<td>0.173</td>
<td>0.161</td>
<td>0.164</td>
<td>0.006</td>
<td>0.089</td>
<td>0.081</td>
<td>0.103</td>
<td>0.093</td>
<td>0.110</td>
<td>0.099</td>
<td>0.093</td>
</tr>
<tr>
<td>G</td>
<td>0.156</td>
<td>0.149</td>
<td>0.174</td>
<td>0.168</td>
<td>0.168</td>
<td>0.097</td>
<td>0.002</td>
<td>0.089</td>
<td>0.114</td>
<td>0.110</td>
<td>0.119</td>
<td>0.110</td>
<td>0.101</td>
</tr>
<tr>
<td>H</td>
<td>0.155</td>
<td>0.149</td>
<td>0.176</td>
<td>0.173</td>
<td>0.163</td>
<td>0.088</td>
<td>0.098</td>
<td>0.001</td>
<td>0.106</td>
<td>0.105</td>
<td>0.126</td>
<td>0.099</td>
<td>0.098</td>
</tr>
<tr>
<td>I</td>
<td>0.156</td>
<td>0.143</td>
<td>0.178</td>
<td>0.170</td>
<td>0.162</td>
<td>0.116</td>
<td>0.128</td>
<td>0.119</td>
<td>0.007</td>
<td>0.079</td>
<td>0.125</td>
<td>0.096</td>
<td>0.089</td>
</tr>
<tr>
<td>J</td>
<td>0.152</td>
<td>0.144</td>
<td>0.181</td>
<td>0.162</td>
<td>0.158</td>
<td>0.102</td>
<td>0.123</td>
<td>0.117</td>
<td>0.085</td>
<td>0.013</td>
<td>0.090</td>
<td>0.083</td>
<td>0.083</td>
</tr>
<tr>
<td>K</td>
<td>0.170</td>
<td>0.161</td>
<td>0.179</td>
<td>0.167</td>
<td>0.165</td>
<td>0.123</td>
<td>0.134</td>
<td>0.144</td>
<td>0.143</td>
<td>0.122</td>
<td>0.004</td>
<td>0.094</td>
<td>0.087</td>
</tr>
<tr>
<td>L</td>
<td>0.148</td>
<td>0.143</td>
<td>0.173</td>
<td>0.165</td>
<td>0.154</td>
<td>0.110</td>
<td>0.124</td>
<td>0.111</td>
<td>0.107</td>
<td>0.098</td>
<td>0.104</td>
<td>0.033</td>
<td>0.051</td>
</tr>
<tr>
<td>M</td>
<td>0.146</td>
<td>0.139</td>
<td>0.168</td>
<td>0.159</td>
<td>0.146</td>
<td>0.103</td>
<td>0.112</td>
<td>0.109</td>
<td>0.097</td>
<td>0.090</td>
<td>0.095</td>
<td>0.054</td>
<td>0.031</td>
</tr>
</tbody>
</table>

Table 2 Partial matrix-correlation tests of geographic distance, historical and contemporary drainage-basin connections, and taxonomic species categories on the haplotype phylogeny.

<table>
<thead>
<tr>
<th>Matrix</th>
<th>$b^*$</th>
<th>$t$-statistic‡</th>
<th>$P$ value‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographic distance</td>
<td>0.157</td>
<td>33.27</td>
<td>0.0001</td>
</tr>
<tr>
<td>Pliocene drainage basins</td>
<td>0.475</td>
<td>95.55</td>
<td>0.0001</td>
</tr>
<tr>
<td>Pleistocene drainage basins</td>
<td>0.237</td>
<td>30.03</td>
<td>0.0001</td>
</tr>
<tr>
<td>Contemporary drainage basins</td>
<td>−0.203</td>
<td>−24.74</td>
<td>0.0001</td>
</tr>
<tr>
<td>Species-level taxonomy</td>
<td>0.084</td>
<td>19.07</td>
<td>0.0011</td>
</tr>
</tbody>
</table>

*Partial regression coefficient; †Student’s $t$; ‡Significance of the partial-regression coefficients and associated $t$-statistics determined by 9999 randomizations of the haplotype patristic-distance matrix.

Table 3 Tajima’s $D$-test for recent population expansion. Estimates of the population parameter $\theta$ ($2N_e\mu$) based on the number of pairwise nucleotide differences ($\theta_S$) and segregating sites ($\theta_S$) for haplotype clades identified in hierarchical phylogenetic analyses (Fig. 4). $n$ is the number individuals sampled in each lineage or subclade. Standard errors of $\theta$ estimates are shown in parentheses. Significantly negative Tajima’s $D$ values (shown in boldface) indicate an excess of young mutations and support the hypothesis of recent population expansion. Lineages F, G, H, and I are excluded from analysis due to the small number of individuals included ($n < 10$).

<table>
<thead>
<tr>
<th>Lineage</th>
<th>$n$</th>
<th>$\theta_S$</th>
<th>$\theta_S$</th>
<th>Tajima’s $D$</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>4.78 (2.88)</td>
<td>4.95 (2.32)</td>
<td>−0.16</td>
<td>0.475</td>
</tr>
<tr>
<td>B</td>
<td>42</td>
<td>1.03 (0.78)</td>
<td>4.18 (1.52)</td>
<td>−2.43</td>
<td>0.001</td>
</tr>
<tr>
<td>C</td>
<td>35</td>
<td>33.70 (16.77)</td>
<td>30.60 (9.39)</td>
<td>−0.34</td>
<td>0.698</td>
</tr>
<tr>
<td>D</td>
<td>48</td>
<td>6.29 (3.27)</td>
<td>11.72 (3.59)</td>
<td>−1.61</td>
<td>0.025</td>
</tr>
<tr>
<td>E</td>
<td>64</td>
<td>27.14 (13.35)</td>
<td>33.20 (9.02)</td>
<td>−0.67</td>
<td>0.288</td>
</tr>
<tr>
<td>J</td>
<td>58</td>
<td>13.54 (6.54)</td>
<td>22.68 (6.28)</td>
<td>−1.47</td>
<td>0.041</td>
</tr>
<tr>
<td>L</td>
<td>60</td>
<td>35.80 (17.52)</td>
<td>41.17 (11.23)</td>
<td>−0.46</td>
<td>0.364</td>
</tr>
<tr>
<td>M</td>
<td>90</td>
<td>32.03 (15.63)</td>
<td>35.89 (9.14)</td>
<td>−0.38</td>
<td>0.391</td>
</tr>
</tbody>
</table>

© 2006 Blackwell Publishing Ltd, Molecular Ecology, 15, 191–207
NCA of lineages E and J reveals phylogeographic patterns consistent with the predictions of the stream-capture hypothesis. The interior clade within a portion of lineage E that spans the Ohio River basin–Atlantic slope divide (subclade 5-1) has a significantly restricted geographic distribution, indicating a contiguous range expansion from the Upper New River drainage system onto the Atlantic Slope (Fig. 5). Similarly, for lineage J, a significantly small interior-tip clade distance suggests that the oldest event shaping geographic patterns of genetic variation in this lineage was a contiguous range expansion onto the Atlantic Coastal Plain from the headwaters of the Nolichucky and Upper Catawba drainage systems (Fig. 6). Fragmentation is inferred between populations in the Neuse + Cape Fear drainages and the remaining samples, which are distributed across tributaries of the Santee, Pee Dee, and Upper Tennessee drainage systems. The genetic signal of range expansion and long-distance colonization from the headwaters of the Upper Tennessee and Upper Catawba is also reflected in less inclusive nested groups of haplotypes (4–8 and 3–20).

The MK test fails to reject the hypothesis of selective neutrality for the ND2 gene, suggesting that our inferences of population history and structure are not likely to be confounded by natural selection acting on molecular variation.

Discussion

Long-term persistence of lineages and population structure

Populations of the *Eurycea bislineata* complex are distributed across all of eastern North America’s major drainage basins, including those in the Ohio River basin and northeast that were strongly affected by Pleistocene glacial advances. Our phylogeographic analysis identifies 13 geographically circumscribed population-level lineages in the group. Molecular divergences between sister lineages indicate that the vast majority of lineages originated through ancient population fragmentation events whose genetic signatures were not erased by Pleistocene drainage modifications (Hocutt et al. 1986; Mayden 1988) and climatic changes (Watts 1980; Webb & Bartlein 1992; Webb et al. 1995). Geographically extensive range expansions are detected for the two lineages (B and D) closest to the receding Wisconsin Ice Sheet, whereas lineages to the south have more restricted geographic ranges that show evidence of long-term population stability (Fig. 2; Table 2). This spatial and temporal pattern of genetic variation matches the leading-edge model of population expansion where lineages that are closest to receding glaciers undergo exponential population...
growth as they colonize recently created habitats, and lineages to the south form barriers to the geographic expansion of neighbouring lineages due to demographic and competitive effects, permitting long-term persistence of their population-genetic structure (Hewitt 1996; Zamudio & Savage 2003). These results, coupled with the extremely low vagility of two-lined salamanders, suggest that many of the lineages have sufficient geographic stability to preserve the phylogenetic signature of ancient hydrological connections and the geographic context of lineage splitting.

**Drainage-basin history and phylogeography**

Geomorphological evidence demonstrates that pre-Pleistocene drainage patterns in eastern North America differed drastically from current hydrological connections. Because of dispersal following the fusion of previously isolated drainages, some lineages of the *E. bislineata* complex no longer exhibit endemism in drainage basins. However, a strong association between historical drainage connections and phylogeographic divergence provides compelling support for the hypothesis that drainage evolution shaped geographic patterns of genetic diversity across eastern North America (Table 2).

A striking feature of the *E. bislineata* complex’s evolutionary history is the strong congruence of phylogeographic patterns with historical, rather than contemporary drainage patterns. The deep phylogeographic split between *E. bislineata*-complex haplotypes sampled from the Lower Tennessee + Cumberland and the Upper Tennessee River drainages is predicted by the hypothesis that the modern Tennessee River drainage formed by fusion of separate palaeodrainage basins that had independent outlets to the Gulf of Mexico until the close of the Pliocene (Figs 1, 2 and 4). Further support for historically separate outlets of the Upper and Lower Tennessee Rivers comes from the phylogenetic relationships of lineages G, H, L and M (Fig. 4). Lineage M contains primarily haplotypes from Upper Tennessee River drainages and is phylogenetically nested inside a clade with an inferred origin in the Gulf drainage basin. Similarly, *Eurycea junaluska* (lineage G: Upper Tennessee) and *Eurycea aquatica* (lineage H: Mobile basin) exhibit a sister-group relationship. The estimated times of divergence between lineages L and M (5%; ∼4 Myr) and G and H (9%; ∼7 Myr) are temporally
congruent with a Miocene to Pliocene formation of the modern Tennessee River basin. Together, these phylogeographic patterns provide strong support for vicariance associated with fragmentation of an ancestral drainage system connecting the Upper Tennessee River and Mobile basin.

Discovery of a deep phylogeographic split separating two distinct evolutionary lineages in the modern Ohio River basin also shows an influence of palaeodrainage basins on current geographic patterns of genetic diversity. The geographic transition between lineages D and E is broadly concordant with the pre-Pleistocene boundary between the Old-Ohio and Teays systems (Figs 1 and 2), suggesting that lineage splitting accompanied geographic isolation imposed by the divide separating these palaeodrainage basins. Furthermore, the divergence for these lineages appears to pre-date the Pleistocene integration of the modern Ohio River basin (6.5% ± 5 Myr). The geographic origins of these lineages, however, require further investigation. Although lineage E haplotypes are clearly associated with drainages of the former Teays River basin, the ancestral location of lineage D haplotypes is ambiguous. Additional sampling and coalescent simulations are needed to determine whether the latter lineage expanded from an Old Ohio, or Teays refugium (e.g. Carstens et al. 2005). Regardless of the geographic context of lineage splitting, these phylogeographic patterns contrast with those of codistributed stream fishes where pre-Pleistocene genetic structure was erased by glaciation and followed by rapid population expansion across most of the modern Ohio River basin (Near et al. 2001; Berendzen et al. 2003).

Secondary contacts between deeply divergent mitochondrial lineages coincide with geographic locations of major drainage divides, a pattern that reinforces the role of drainage basins in generating contemporary geographic patterns of genetic variation. For example, sympathy between lineages that span the basin north–south split in the phylogeny closely parallels the juxtaposition of the New, Tennessee, and Atlantic Slope divides in the southern Appalachian Highlands. Additional such zones of sympathy between lineages occur along the Ohio–Cumberland and Tennessee River–Gulf divides.

Despite its potential as a contemporary isolating mechanism, the ECD has migrated westward as Atlantic Slope drainage systems eroded through the Blue Ridge, fusing with and reversing the flow of western headwater drainage systems. Geological evidence indicates headwater capture of the New by the James and Roanoke Rivers (Ross 1969; Hocutt et al. 1986), and tributaries of the Upper Tennessee by the Santee and Savannah River drainages (Ross 1971; Starnes & Etnier 1986). Freshwater-fish distributions also indicate multiple faunal invasions of the Atlantic Slope from drainages west of the ECD facilitated by these stream captures (Hocutt et al. 1986; Starnes & Etnier 1986).

Geographic genetic variation in the *E. bislineata* complex supports a role for colonization of the Atlantic Slope associated with headwater stream capture. Haplotypes sampled from the Kanawha–New River drainage span the deepest phylogenetic divergence within lineage E, suggesting a Teays/Ohio River basin origin for this lineage with subsequent dispersal to the James, Roanoke, and Santee drainages on the Atlantic Slope, and Holston drainage in the Upper Tennessee River basin. Geological evidence also supports the New–Roanoke–James and New–Holston exchanges (Ross 1969; Hocutt et al. 1986). The haplotype phylogeny identifies two dispersal events to the Atlantic Slope in the vicinity of the drainage divide separating the Kanawha + New and Roanoke + James systems. Extensive geographic sampling for nested group 6-1 of lineage E reveals a range expansion from the headwaters of the New and Roanoke Rivers into the James River basin, the genetic pattern predicted by an eastward dispersal event facilitated by a series of hypothesized Pleistocene stream captures between these drainage systems. A recent fish phylogeographic study also demonstrates similar dispersal between the New and Roanoke drainage systems (Berendzen et al. 2003), suggesting that stream capture played a key role in shaping current geographic patterns of genetic and species diversity along the Atlantic Slope.

Within the southern clade, lineage J exhibits recent range expansion from the Nolichucky (Upper Tennessee) and headwaters of the upper Catawba (Santee) onto the Atlantic Coastal Plain. This phylogeographic pattern is consistent with the hypothesis that population expansion across the ECD was associated with a geologically inferred stream capture between these drainage systems (Ross 1971), and it provides strong support for the hypothesis that piracies of the upper Tennessee by headwaters of the Santee drainage system created dispersal corridors across the ECD for stream-dependent faunas (Starnes & Etnier 1986).

**Gene trees and species limits**

An important strength of using mtDNA haplotype phylogenies to diagnose independent evolutionary lineages is that mitochondrial genes have an effective population size one quarter that of nuclear genes (Wiens & Penkrot 2002). As a result, species become reciprocally monophyletic for mtDNA haplotypes more quickly than for those of neutral nuclear genes (Moore 1995; but see Hoelzer 1997). The ancient history of phylogeographic fragmentation uncovered within the *E. bislineata* complex and its strong congruence with historical drainage connections suggest that the current taxonomy underestimates species-level diversity of the group and obscures its biogeographic history. Moreover, few haplotypes are shared among sampling locations, indicating that even at local spatial scales, *E. bislineata*-complex populations are highly genetically structured, a
result suggesting that ongoing gene flow is not likely to be a homogenizing force where the geographic ranges of major population-level lineages abut one another.

_Eurycea aquatica_ and _E. junaquiska_ each form monophyletic groups of haplotypes and are morphologically diagnosable from other sympatric lineages of the complex (Rose & Bush 1963; Sever _et al._ 1976; Ryan 1997). Thus, in contrast with earlier work that synonymized these taxa with _Eurycea cirrigera_ (Jacobs 1987), our study provides strong support for their recognition as species under the phylogenetic, evolutionary, and biological species concepts (de Queiroz 1998).

_Eurycea bislineata_, _E. cirrigera_, and _Eurycea wilderae_ each include deeply divergent haplotype clades that replace each other geographically. The distributions of lineages A, B, D, E, L, and M are broadly congruent with geographic limits of cryptic, allozyme-based genetic groupings (Jacobs 1987), supporting the hypothesis that these haplotype clades diagnose cryptic, independent evolutionary lineages. Moreover, a combination of genetic, ecological, and sexual-isolation analyses across contact zones between lineages L and M (Camp _et al._ 2000; Kozak & Montanucci 2001; Kozak 2003) and B and D (Guttman & Karlin 1986) provide independent evidence for a lack of genetic and/or ecological exchangeability (Templeton 2001) between these deeply divergent, parapatric haplotype clades. Lineages C, F, I, and K lack independent confirmation of their phylogenetic distinctness.

One potential drawback of using mtDNA to delimit species is that it may be particularly susceptible to introgression across species boundaries, leading to discordance between gene trees and species trees (Good _et al._ 2003; Glor _et al._ 2004; Morando _et al._ 2004; Weisrock _et al._ 2005). The presence of haplotypes from both _E. cirrigera_ and _E. wilderae_ in lineages E and F in a region of sympathy with lineage M in the southern Appalachian Highlands may indicate mtDNA introgression across the boundaries of distinct evolutionary lineages. Such geographically structured discordance is not expected for retention of ancestral polymorphisms (Good _et al._ 2003). Furthermore, the close association of these lineages with the New, Tennessee, and Atlantic Slope drainage divides is suggestive of hybridization along the margins of historical barriers to gene exchange. Alternatively, the colour morphology used to diagnose _E. wilderae_ may have evolved repeatedly during the evolutionary history of the _E. bislineata_ complex, with the haplotype phylogeny providing a more accurate picture of the geographic limits of independent evolutionary lineages. Supporting the latter hypothesis, geographic patterns of genetic variation in allozymes also demonstrate that _E. wilderae_ is nonmonophyletic and reveal an ancient genetic break separating populations in the New and Tennessee River drainages (Jacobs 1987). Further fine-scale population-genetic studies that employ multiple unlinked molecular markers are needed to discriminate these alternatives.

Conclusions

The geological history of stream drainage basins provides a wealth of a priori hypotheses for the geographic fragmentation of populations of stream-dependent organisms that are ideally suited for testing with phylogeographic methodologies. Here, we demonstrate that salamanders of the _E. bislineata_ complex exhibit a complex phylogeographic history that preserves the genetic signatures of palaeodrainage connections that were altered by glacial advances and headwater erosion. Our molecular genetic results reinforce geological and faunistic evidence for composite origins of major eastern North American drainage systems. More generally, our study provides a general methodological framework for combining independent biological and geological information to formulate and to test hypotheses of the complex interaction between historical and contemporary factors that generate and maintain patterns of biodiversity in stream-dependent organisms. The generality of the phylogeographic hypotheses tested in this study should be investigated using comparable analyses in other codistributed stream-dependent species.

Acknowledgements

We thank R. Bonett, C. Camp, R. Highton, D. Huéstis, J. Marshall, K. Zamudio, W. Savage, W. Van Devender, and the LSU Collection of Genetic Resources for donating tissue samples for genetic analysis. R. Bonett, M. Gifford, and three anonymous reviewers provided helpful comments on the manuscript. The American Society of Ichthyologists and Herpetologists’ Gaige Fund, Highlands Biological Station’s Bruce Family and Charles Ash Scholarships, and a National Science Foundation Doctoral Dissertation Improvement Grant DEB-0206739 generously supported this research.

Supplementary material

The supplementary material is available from http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC2757/MEC2757sm.htm

Table S1 Geographic sampling information for the salamanders used in this study. Population numbers correspond to Fig. 2.

Table S2 List of _Eurycea bislineata_ complex haplotypes sequenced in this study, the localities in which they were found, field numbers of specimens from which they were sampled, and GenBank Accession numbers. Haplotypes are named according to the taxonomic species from which they were sampled. Populations are numbered as in Table S1, and Figs 2 and 4, with the number of individuals sampled for each haplotype shown in parentheses. KHK, Kenneth H. Kozak field series; RMB, Ronald M. Bonett field series; WKS, Wesley K. Savage field series; H, Louisiana State University Museum of Natural History Collection of Genetic Resources; numbers without a prefix indicate Richard Highton field numbers. Voucher specimens from KHK’s field series will be deposited in the LSU Museum of Natural History following a formal taxonomic revision of the complex.

© 2006 Blackwell Publishing Ltd, Molecular Ecology, 15, 191–207
References


Bremner K (1994) Branch support and tree stability. Cladistics, 10, 295–304.


Maddison WP, Maddison DR (1992) MACCLADE: analysis of phylogeny


This study was conducted as part of Kenneth Kozak’s PhD thesis focused on investigating patterns of speciation, ecological diversification, and community assembly in eastern North American plethodontid salamanders. He is currently a National Science Foundation Postdoctoral Fellow in Bioinformatics at Stony Brook University where he is investigating the relationship between landscape ecology and spatial patterns of genetic variation and species richness in woodland salamanders. Russell Blaine is a PhD candidate at Washington University. His research is focused on comparative phylogeography and landscape genetics of reptiles in the desert southwest of North America. Allan Larson is Professor of Biology at Washington University. His research specialty is molecular phylogenetics and evolution of vertebrates, particularly lizards and salamanders.