Nested Cladistic Analysis Indicates Population Fragmentation Shapes Genetic Diversity in a Freshwater Mussel

Thomas F. Turner,* Joel C. Trexler,† John L. Harris‡ and Jody L. Haynes§

**Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico 87131, †Department of Biological Sciences, Florida International University, Miami, Florida 33199, ‡12301 Pleasant Forest Drive, Little Rock, Arkansas 72212 and §Tropical Research and Education Center, University of Florida, Homestead, Florida 33031*

Manuscript received February 24, 1999
Accepted for publication October 4, 1999

ABSTRACT

Recently developed phylogeographic analyses that incorporate genealogical relationships of alleles offer the exciting prospect of disentangling historical from contemporary events. However, the relative advantages and shortfalls of this approach remain to be studied. We compared the nested cladistic method to the more traditional analysis of variance approach in a study of intraspecific genetic variation in the freshwater mussel, *Lampsilis hydiana*. We surveyed 257 specimens for nucleotide sequence level variation in a fragment of the mitochondrial 16S rRNA gene. When compared side by side, nested cladistic analysis and analysis of molecular variance (AMOVA) identified fragmentation of Arkansas river populations from remaining populations to the southwest. Nested cladistic analysis identified a second, more recent separation of Ouachita and Upper Saline river populations that was not detected by AMOVA. Differences among analytical methods probably arise from treatment of spatial hierarchical information: hierarchical groups emerge via a parsimony criterion in nested cladistic analysis but must be specified *a priori* in AMOVA. Both methods identified significant genetic structure among localities within hierarchical groups. Results from AMOVA suggested little gene flow among local populations with an island model. However, inferences about process that gave rise to patterns at this level were not possible in nested cladistic analysis, because an ancestral (interior) haplotype was not observed for a key one-step clade in the parsimony network. Our results suggest that, under some circumstances, nested cladistic analysis has lower power than more traditional analysis of variance to infer processes at the local population level.

ONE of the most refractory, yet fundamental, areas of inquiry in evolutionary biology is partitioning the effects of historical and contemporary processes in shaping geographic patterns of genetic diversity (Avise *et al.* 1987). Standard approaches to describing population genetic structure and estimating gene flow based on Wright's island model, while powerful, do not attempt to disentangle past events from contemporary processes (Neigel 1997). Molecular genetic data provide the exciting prospect of recording genealogical relationships in a manner that can be used to disentangle past and present processes, though analytical tools to procure this insight are not well developed. Templeton *et al.* (1995) and Templeton (1998) have advanced this new class of phylogenetically based gene flow analysis with their nested cladistic approach. We have used this approach to analyze patterns of genetic variation in unionid mussels (Mollusca: Bilvalvia: Unionacea) that inhabit mountain streams in the central highlands of the United States. The nested cladistic approach potentially provides rich insight into the relative roles of biogeographic and current ecological processes in shaping genetic diversity observed in our data.

Geographic patterns of genetic diversity of unionids are likely to reflect a substantial historical element because of their apparently sedentary lifestyle, specialized habitat requirements, and geographic distribution. These bivalves are very diverse in the Mississippi basin, (van der Schalie and van der Schalie 1950; Burch 1975), an ancient basin known to encompass significant biogeographic subdivision in other taxa, *e.g.*, fishes (Mayden 1988). The drainage basin crosses from northern North American regions that experienced Pleistocene glaciation to southern regions not covered by glaciers at that time. This history has left marked effects on the mussel fauna (Johnson 1980). The Unionidae includes many species with restricted ranges limited to small systems of interconnected drainage basins. The potential for local isolation is probably heavily influenced by the larval stage, which is an obligate parasite of fish hosts. This phase is thought to be the major source of dispersal for these mussels (Pennack 1978). It is likely that recurrent gene flow is limited by the diversity of fish hosts and the dispersal patterns of those hosts. Since many stream fishes have limited ability to pass through high order streams and rivers, mussels should be similarly limited. We have examined mito-

Corresponding author: Thomas F. Turner, Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM 87131. E-mail: turnert@unm.edu

chondrial DNA sequences from freshwater mussels from the central highlands of Arkansas to estimate gene flow and historical, biogeographic patterns in their genetic diversity. These data provide an excellent opportunity to compare inferences about the processes that shape genetic diversity between nonhistorical analysis of molecular variance (AMOVA) and the nested cladistic method of Templeton *et al.* (1995).

MATERIALS AND METHODS

A total of 257 *Lampsilis hydiana* were collected between 1991 and 1993 from 17 localities distributed over five major river drainages in southwestern Arkansas and east Texas (Figure 1; Table 1). The study area covers roughly one-third the geographic range of *L. hydiana.* In addition, specimens of *L. cardium, L. powelli, L. reeveiana, L. satura, L. siliquoidea,* and *L. teres* were collected from various rivers and streams throughout the study area (Figure 1; Table 1). These species co-occur with *L. hydiana* over some or all of its geographic range (Harris and Gordon 1990). *L. powelli,* a federally threatened species, was sampled nondestructively by clipping a small piece of muscular foot tissue. For all other taxa, entire soft tissues were removed and stored immediately in liquid nitrogen.

Genomic DNA was isolated from 20–200 mg of freeze-dried foot tissue using the method of Tyler *et al.* (1995), omitting the RNAse step. A portion of the 16S mitochondrial ribosomal RNA gene (16S rRNA) was amplified via polymerase chain reaction (PCR) using primers 16Sint3-L 5′-TGAGCGTVCTAA GGTAGC-3′ and 16Sint4-H 5′-AKCCAACATCGAGGTCG CAA-3′. PCR was conducted in 7-μl volumes, containing 1 μl (~100 ng) of sample DNA, 0.7 μl 10× reaction buffer, 200 μm each dGTP, dATP, dTTP, dCTP, 1.5 mm $MgCl_2$, and 0.4 units of *Thermus aquaticus* (*Taq*) DNA polymerase. Primers were added to the reaction at a final concentration of 0.8 μm, which included 0.1 μm of each primer end-labeled with [γ^{32}]ATP using T4 polynucleotide kinase. PCR consisted of 30 cycles of denaturation at 94° for 30 sec, annealing at 50° for 30 sec, and extension at 72° for 1 min.

Individual variation was screened by examining single-stranded conformation polymorphism (SSCP) of 16S rRNA fragments. PCR amplification products were diluted 1:5 in denaturing formamide buffer (Sambrook *et al.* 1989), heated to 95° for 5 min, and then immediately chilled in an ice slurry. Samples were loaded on 5% nondenaturing polyacrylamide (37.5:1 acrylamide:bis-acrylamide) gels with 5% glycerol and subjected to electrophoresis at 10 W for 15 hr at room temperature (25–27°) in 1× TBE buffer. Single-stranded conformations of 16S rRNA fragments were visualized by autoradiography and scored by comparing relative mobilities.

An interpretation of patterns of mtDNA variation that assumes strict maternal inheritance may be complicated by the presence of paternal haplotypes that are transmitted to male offspring, a phenomenon termed doubly uniparental inheritance, or DUI (Skibinski *et al.* 1994; Zouros *et al.* 1994). DUI has been observed in another unionid species, *Anodonta grandis* (Liu *et al.* 1996), and may be present in genus Lampsilis. We minimized the possibility of analyzing male mtDNA haplotypes by restricting the source for DNA isolation to foot tissue for all specimens. In marine mussels genus Mytilus, male mtDNA haplotypes are dominant in male gonads, whereas female mtDNA haplotypes are dominant in somatic tissues (Garrido-Ramos *et al.* 1998). As an additional precaution, SSCP electromorphs were scrutinized for the presence of four bands, which would indicate the presence of an additional mtDNA haplotype.

At least two representatives of each SSCP variant (haplotype) were sequenced directly using the Amplitaq cycle sequencing kit and following instructions supplied by the manufacturer (Perkin Elmer, Branchburg, NJ). Cycle sequencing was done in both directions (90% overlap) using primers listed above. Sequences were aligned using the PILEUP module of GCG (Devereux *et al.* 1984) and adjusted by eye. Additional 16S rRNA data from a previously published study of freshwater mussels (Lydeard *et al.* 1996) were included in our study. Sequences from the following species (with GenBank accession nos.) were selected: *L. teres* (U72568), *L. siliquoidea* (U72571), *Villosa delumbis (*U72574), *Medionidus conradicus* (U72572), *Elliptio dilatata* (U72557), *Amblema plicata* (U72548), *Quadrula quadrula* (U72552), *Anodonta couperiana* (U72560), *Margaritifera margaritifera* (U72544). These sequences were aligned with our data by eye and phylogenetic analysis was performed on the entire data set [21 operational taxonomic units (OTUs)].

Phylogenetic analysis of 16S ribosomal gene sequences was conducted using maximum parsimony (MP) and full heuristic searches in PAUP* (version 4.1; Swofford 1995). All searches employed the following options: characters were unordered, uninformative characters were ignored, taxa were added to starting trees at random with 10 replications per run, multistate taxa were treated as uncertain, minimal trees were kept, tree swapping was conducted by tree bisection and reconnection, zero length branches were collapsed, and insertions and deletions (indels) were treated as a "fifth base." Each character set was resampled with replacement producing 200 bootstrap replicates subjected to full heuristic searches as described above. The proportion of bootstrap trees that resolved a particular branch was interpreted as an index of support for that

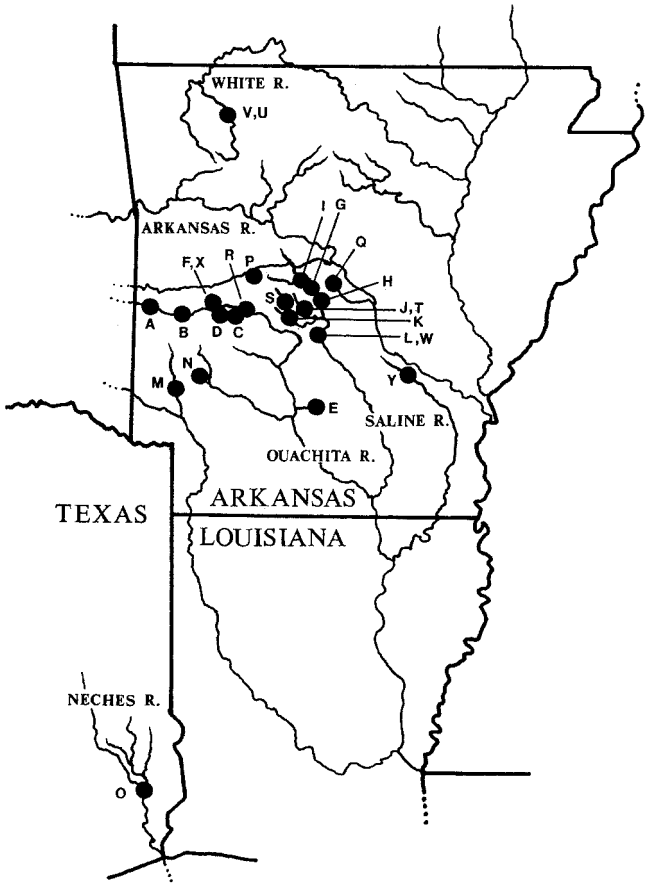

Figure 1.—Sampling localities of *Lampsilis* sp. surveyed for variation at the mitochondrial 16S rRNA gene. Map symbols correspond to locality information in Table 1.

TABLE 1

Sample sizes and collection localities of freshwater mussels (genus Lampsilis, family Unionidae) examined in this study

Species	Pop.	Map symbol	*N*	Collection locality	Latitude, longitude
L. hydiana	LH 1	A	12	Lewis Creek, Ouachita R., AR	34° 38′ N, 94° 13′ W
	LH 2	B	19	Cherry Hill, Ouachita R., AR	34° 34′ N, 94° 00′ W
	LH 3	C	17	South Fork Ouachita R., AR	34° 34′ N, 93° 43′ W
	LH 4	D	6	South Fork Ouachita R., AR	34° 34′ N, 93° 44′ W
	LH 5	E	12	White Oak Creek, Ouachita R., AR	33° 52′ N, 92° 45′ W
	LH 6	F	5	Oden, Ouachita River, AR	34° 37′ N, 93° 47′ W
	LH 7	G	16	N. Fork Saline R., AR	34° 42′ N, 92° 39′ W
	LH 8	H	10	N. Fork Saline R., AR	34° 39′ N, 92° 37′ W
	LH 9	I	22	N. Fork Saline R., AR	34° 46′ N, 92° 44′ W
	LH 10	J	2	Alum Fork Saline R., AR	34° 47′ N, 92° 53′ W
	LH 11	K	2	S. Fork Saline R., AR	34° 33′ N, 92° 46′ W
	LH 12	L	22	Drop Creek, Saline R., AR	34° 27′ N, 92° 37′ W
	LH 13	M	14	Red R. drainage, AR	33° 58′ N, 94° 04′ W
	LH 14	N	22	Muddy Fork Little Missouri R., AR	34° 05′ N, 93° 45′ W
	LH 15	O	15	Neches R., TX	30° 47′ N, 94° 09′ W
L. cf. hydiana	LcfH 16	P	18	S. Fork Fourche la Fave R., AR	34° 53′ N, 93° 05′ W
	LcfH 17	Q	25	Fourche Creek, Arkansas R., AR	34° 43′ N, 92° 31′ W
L. powelli	LP 1	R	12	S. Fork Ouachita R., AR	34° 34′ N, 93° 38′ W
	LP 2	S	13	Middle Fork Saline R., AR	34° 40′ N, 92° 53′ W
	LP 3	T	15	Alum Fork Saline R., AR	34° 41′ N, 92° 51′ W
L. siliquoidea	LSi 1	U	17	War Eagle Creek, White R., AR	36° 12′ N, 93° 51′ W
L. reeveiana	LR 1	V	11	War Eagle Creek, White R., AR	36° 12′ N, 93° 51′ W
L. cardium	LC 1	W	3	Drop Creek, Saline R., AR	34° 27′ N, 92° 37′ W
L. satura	LS 1	X	11	S. Fork Fourche la Fave R., AR	34° 53′ N, 93° 05′ W
L. teres	LT 1	Y	6	Bayou Bartholomew R., AR	34° 11′ N, 91° 57′ W

Map symbols correspond to Figure 1.

branch. The next step was to determine a nesting design based on the evolutionary relationships of haplotypes for nested cladistic analysis of *L. hydiana*. We began by asking whether maximum parsimony was justified to reconstruct intraspecific relationships of haplotypes by computing the probability that a site difference between two randomly drawn haplotypes resulted from more than one mutational event (the nonparsimonious state; Hudson 1989). The parameter θ was estimated for haplotypes identified from *L. hydiana* using equation 10.5 in Nei (1987, p. 256). Theta is equal to $4N\mu$, where N is the effective population size and μ is the mutation rate. The estimate of θ was substituted into equation 10 of Hudson (1989) and solved using a program written by A. R. Templeton for the MATHEMATICA computer package.

The intraspecific maximum parsimony network defined an evolutionary hierarchy of haplotypes (zero step clades) nested within one-step clades (haplotypes that differed by one mutational step grouped together), nested within two-step clades (haplotypes that differed by two mutational steps grouped together). Clade distances (D_c), or the geographical ranges of clades at each hierarchical level, were estimated by calculating the average distance that individuals bearing haplotypes grouped in a particular clade deviated from the geographic center of that clade (Templeton *et al.* 1995). Nested clade distances (D_n) were determined by calculating the average distance of observations falling within a particular haplotype class from the geographic center of the entire nesting clade. The difference of interior (ancestral) and tip clade D_c and D_n distances were also calculated when an interior clade existed to yield $D_cI - D_cT$ and $D_nI - D_nT$ values, respectively. Methods are discussed in more detail in Templeton *et al.* (1995) and Templeton (1998). Our analysis differs from Templeton *et al.* (1995) in one important way; rather than use great circle distances we used distance in river kilometers between sampling localities to estimate D_c and D_n values.

The null hypothesis of no geographical associations of clades and nesting clades was tested in two ways. First, sampling locations were treated as categorical variables, and homogeneity tests were employed to ask whether clades at each nesting level were distributed at random across sampling locations. For each level of the nesting design, significant nonrandom associations of sampling locations and clades were identified by comparing observed chi-square values to a distribution of chi-square values generated from 1000 permutations of the original data (Roff and Bentzen 1989). A second method incorporating D_c and D_n was employed to account for geographic information. The hypothesis that dispersion distance of clades was not greater or less than expected at random was tested by comparing observed D_c and D_n values to a distribution of D_c and D_n calculated for each of 1000 random permutations of clades against sampling locations (Templeton *et al.* 1995). Permutation tests were conducted separately for each level of the nested cladogram. Analyses were conducted using the GEODIS version 2.0 computer package (Posada *et al.* 1999). Once significance levels for D_c and D_n were determined, inferences about processes that likely gave rise to patterns of clade dispersion were made using the table presented in the appendix of Templeton *et al.* (1995).

In addition to nested cladistic analysis, we used AMOVA to partition genetic variation into components attributable to differences among localities within prespecified hierarchical groups (Φ_{SC}), among localities across the entire study area (Φ_{ST}), and among prespecified hierarchical groups (Φ_{CT}; Excoffier *et al.* 1992). Two hierarchies were analyzed. First,

TABLE 2

16S rRNA-SSCP haplotype frequencies by collection locality

	Haplotype											
Locality	1	2	3	4	5	6	7	8	9	10	11	12
LH 1	28		2									
LH 2	10	1	8									
LH 3	15	2										
LH 4	5	1										
LH 5	7	4	1									
LH 6		2	3									
LH 7		3	13									
LH 8	1	1	8									
LH 9	4		18									
LH 10	1		1									
LH 11			2									
LH 12	19		3									
LH 13	14											
LH 14	11	2	5	4								
LH 15	4	6	5									
LcfH 16					18							
LcfH 17					25							
LP 1							12					
LP 2			1				12					
LP 3			1			2	13					
LSi 1							17					
LR 1								2	9			
LC 1										3		
LS 1										4	7	
LT 1												6

Locality numbers LH1 through LcfH17 (*L. hydiana*) were included in AMOVA and phylogeographic analysis.

two groups of localities were distinguished based on external (shell) morphological differences: (1) Arkansas River localities numbered LcfH16 and LcfH17, and (2) all other sample drainages numbered LH1 through LH15. Second, three groups of localities were distinguished based on sample river drainage excluding localities LcfH16 and LcfH17 in the Arkansas River. Groups were as follows: (1) Ouachita River, LH1 through LH6, LH13, LH14; (2) Saline River, LH7 through LH12; and (3) Neches River, LH15.

We also tested for isolation-by-distance among localities LH1 through LH15 using the method of Rousset (1996). For each pair of localities, we calculated the quantity $\Phi/(1-\Phi)$ and the river distance (kilometers) separating each pair of localities. This procedure resulted in 105 pairwise comparisons. Pairwise values of $\Phi/(1-\Phi)$ were regressed on pairwise river distance using the ordinary least-squares method. A slope that differed significantly from zero (via *t*-test with d.f. = 1, 13 adjusted to reflect only the number of original localities not the total number of pairwise comparisons—see Hellberg 1996) was interpreted as evidence for isolation-by-distance over this portion of the species range (Rousset 1996). Analyses of molecular variance were conducted using the WINAMOVA computer package (Excoffier *et al.* 1992).

RESULTS

A total of 12 unique haplotypes were identified using SSCP analysis of 16S rRNA fragments (Table 2). No more than two bands were detected for any individual assayed, suggesting that fragments analyzed were a single (probably female) mtDNA lineage. Nucleotide sequencing revealed that SSCP variation among haplotypes resulted mostly from nucleotide substitutions and indels. Indels at positions 234, 235, and 236 were parsimony informative, but the remainder were autapomorphic. Nucleotide sequences of 16S rRNA fragments are deposited in GenBank under accession nos. AF191565–AF191576, and sequence alignment is available from T. F. Turner upon request.

Phylogenetic analysis using maximum parsimony produced a strict consensus tree (of nine equally parsimonious trees) that agreed well with results presented in Lydeard *et al.* (1996). Sequences we obtained for *L. teres* and *L. siliquoidea* clustered with sequences determined for these taxa by Lydeard *et al.* (1996) with 100% bootstrap support (Figure 2). In general, bootstrap support for reciprocal monophyly of each species was high ($\geq$80%), but deeper relationships within genus Lampsilis were not as strongly supported. The strict consensus tree indicated that *L. hydiana* and *L. powelli* plus *L. siliquoidea* were monophyletic and *L. reeveiana*, *L. cardium* plus *L. satura*, and *L. teres* were increasingly distantly related to this group (Figure 2). Bootstrap support for these relationships was $<$50% in all cases. In addition

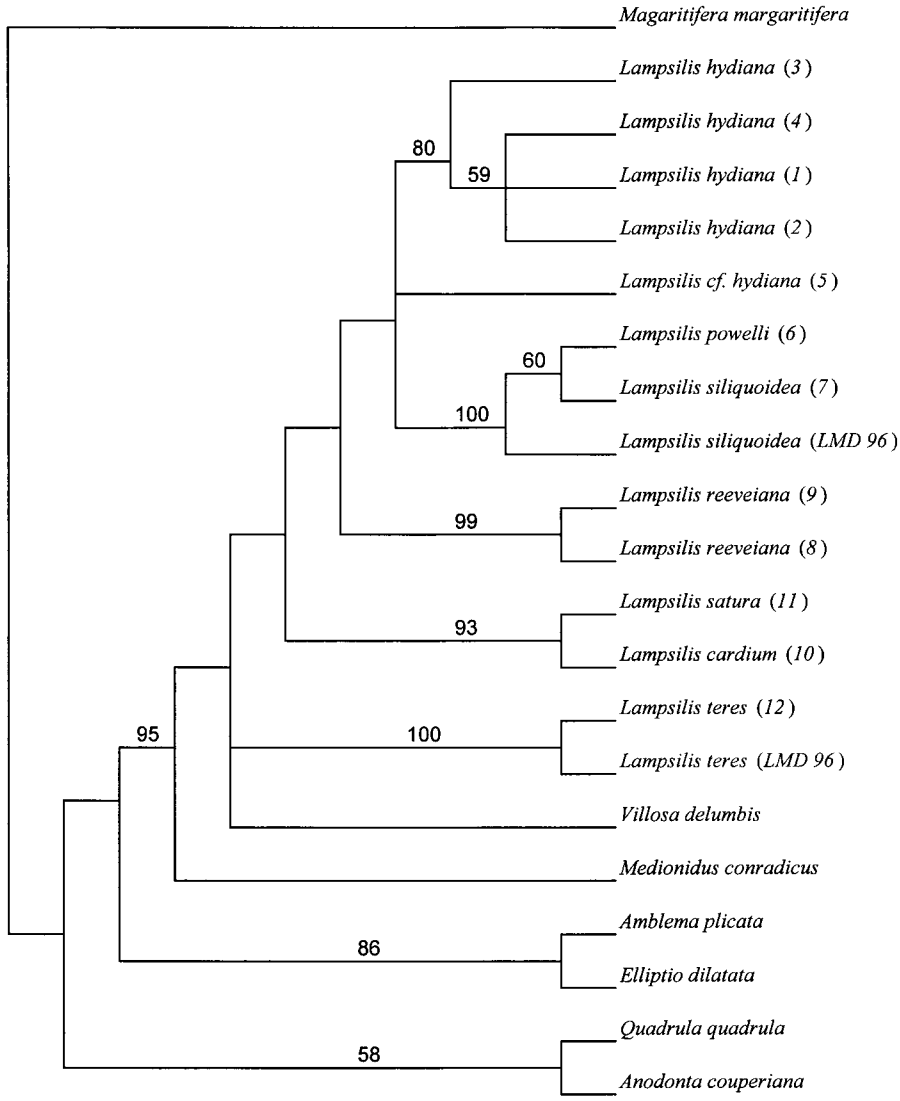

Figure 2.—A strict consensus tree of nine equally parsimonious trees obtained using maximum parsimony analysis of 16S rRNA sequences. In this analysis, insertions and deletions were treated as a fifth base, and all changes received equal weight. The numbers at the nodes indicate the percent of bootstrap trees (of 200 replicates) that supported a particular branch. Branches that received <50% support are not labeled. Numbers in parentheses following species names correspond to haplotype numbers in Table 2. Nucleotide sequences from members of the genus Lampsilis obtained by Lydeard *et al.* (1996) are labeled LMD 96. Sequences for taxa other than for members of the genus Lampsilis were obtained by Lydeard *et al.* (1996) and were used as outgroups to place our study species into phylogenetic context.

to being consistent with Lydeard *et al.* (1996), the topology of the strict consensus tree was also consistent with an unpublished allozyme data set and with interspecific variation in shell morphology.

Maximum parsimony appeared to be justified for reconstructing intraspecific relationships of *L. hydiana* haplotypes. Theta was estimated as $\theta = 0.0074$ for five haplotypes isolated from *L. hydiana*. Theta was substituted into the Hudson (1989) equation, and the solution indicated that for any given site, a difference observed between two randomly drawn haplotypes was unlikely to have arisen from more than one mutational change ($P = 0.034$).

The nesting design that resulted from maximum parsimony analysis indicated that *L. hydiana* haplotypes 1, 2, and 4 appeared at the tips and that haplotype 3 was interior, or ancestral (Figure 3). Together, these four haplotypes defined nesting clade 2-1. Haplotype 5 was fixed in two localities from the Arkansas River basin. Individuals at these localities, designated *L.cf. hydiana*, were morphologically distinct from remaining *L. hydiana* populations to the southwest. Haplotype 5 defined a second nesting clade 2-2, which was interior to nesting clade 2-1. The two nesting clades 2-1 and 2-2 were separated by nine mutational steps (Figure 3).

Homogeneity testing revealed significant nonrandom association of clades and sampling locations, indicating phylogeographic structure in the data at higher clade levels (Table 3). Estimation of D_c and D_n indicated that haplotypes located at tips (haplotypes 1, 2, and 4) varied in their distribution over the study range. Haplotype 4 was restricted to a single locality in the Little Missouri River (Ouachita Drainage) and had significantly small D_c (Figure 4). Clade (D_c) and nested clade (D_n) distances for nesting clade 1-1 were significantly small. The distribution of this nesting clade included all localities in the Ouachita drainage and lower Saline River, but it occurred at very low frequencies in the upper Saline River. Nesting clade 1-2 was significantly more widespread than randomized data owing to the wide distribution of haplotype 2, which was found in low to moderate frequencies in the Ouachita, Saline, and Neches river

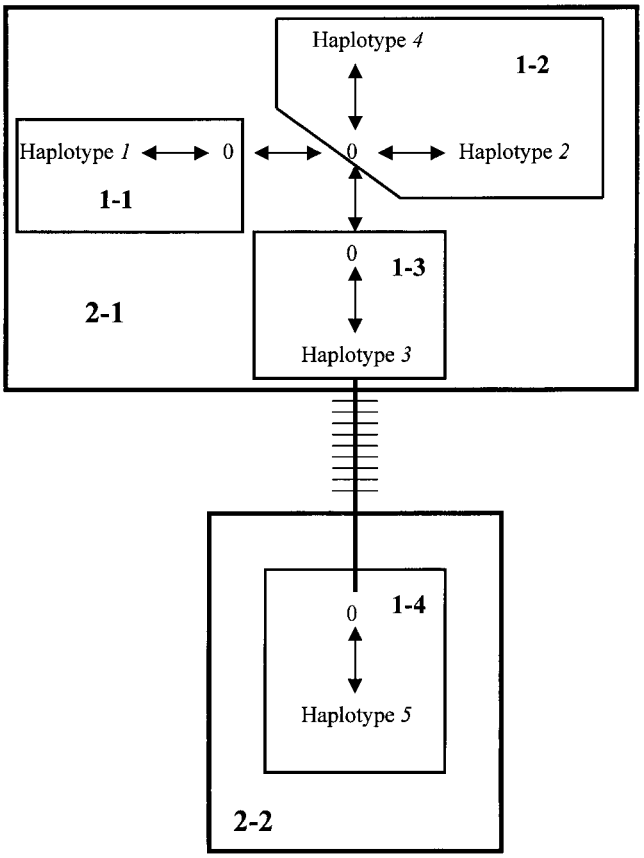

Figure 3.—The nesting design inferred from maximum parsimony for phylogeographic analysis of *L. hydiana.* Haplotype numbers are the same as in Table 2. Each clade and nesting clade has numbers that designate the number of mutational steps in the nesting clade and an identification number, respectively.

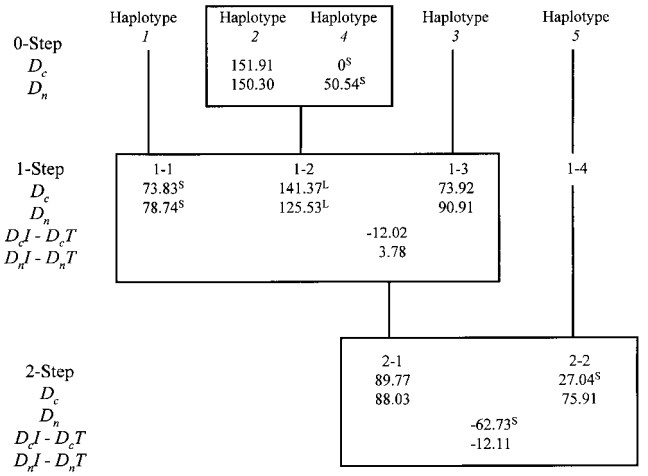

Figure 4.—Results of the nested geographical analysis of *L. hydiana* 16S rRNA haplotypes. Clade (D_c) and nesting clade (D_n) distances are reported for each level of the nesting design and superscripts refer to significantly small (S) or large (L) clade and nested clade distances. The average difference between interior (I) and tip (T) D_c and D_n values is presented in rows labeled I-T.

(TX) drainages. The interior haplotype 3 was also widespread, but exhibited large differences in frequency across its distribution, reaching maximum frequencies in the upper reach of the Saline River. Haplotype 5 was restricted to the Arkansas River drainage. Geographic distributions of clades and nested clades indicated two well-supported population fragmentation events (Table 4).

AMOVA revealed significant spatial partitioning of genetic variation at all hierarchical levels, but the magnitude of Φ-values, and thus the interpretation of patterns of genetic variance, depended on the hierarchical sampling design employed. For example, when morphologically distinct localities in the Arkansas River were grouped and compared to all other localities, the majority of variance was attributable to differences among hierarchical groups ($\Phi_{CT} = 0.84$, $P < 0.001$) and a lower proportion of variation was attributable to differences among localities within groups ($\Phi_{SC} = 0.327$, $P < 0.001$). The second hierarchical design that compared three groups sorted by river drainage (excluding morphologically distinct Arkansas River localities) revealed the converse pattern. In this case, among-group variance was much lower $\Phi_{CT} = 0.188$ ($P < 0.001$) than within-group variance $\Phi_{SC} = 0.431$ ($P < 0.001$), indicating that the majority of genetic variance was explained at the level of localities within, not among, river drainages. The slope of regression of pairwise $\Phi/(1 - \Phi)$ and river distance was not significantly different from zero ($r^2 = 0.019$, $t = 1.44$, $P = 0.17$), indicating that isolation-by-distance played little or no role in structuring genetic variation.

Side-by-side comparison of nested cladistic analysis and analysis of molecular variance indicated congruent support for two of three inferences made about processes influencing spatial genetic variation in *L. hydiana* (Table 5). A fragmentation event separating Arkansas River localities from all other *L. hydiana* localities was supported by both methods. Isolation-by-distance was not supported by either method. However, nested cladistic analysis identified a fragmentation event between upper Saline River and the remainder of *L. hydiana*

TABLE 3

Nested contingency analysis of geographical associations

Clade	Permutational chi-square statistic	Probability
1-2	15.8	0.056
2-1	142.7	<0.001
Entire cladogram	257.0	<0.001

The clade column refers to the nesting clade in Figure 3. Clades with no genetic and/or geographical variation within them are not given as no test is possible within such nested categories.

TABLE 4

Inference chain based on results of geographic dispersion analysis given in Figure 4

Clade	Chain of inference	Inference
Haplotypes nested in 1-2	1,2 (no interior clade)	Significantly small clade distance for haplotype (LH92311)—further inference requires interior clade
One-step clades nested in 2-1	1,2,3,5,15, NO	Past fragmentation
Two-step clades nested in the entire cladogram	Nine mutational steps between 2-1 and 2-2	Allopatric fragmentation

Inference chain is found in the appendix of Templeton *et al.* (1995).

localities (excluding Arkansas River localities) that went undetected by AMOVA (Table 5).

DISCUSSION

Intraspecific patterns of genetic variation can provide important insights into the roles of current and historical processes in species formation (Bernatchez and Wilson 1998) and provide evidence for defining regional boundaries for studies of ecological communities (*sensu* Ricklefs 1987; Gorman 1992). However, any interpretation of this type requires statistically rigorous separation of current and historical processes that shape genetic variation. Analyses of genetic variance based on Wright's island model incorporate spatial patterns of allelic distribution and are limited to the interpretation that such patterns arise from the interplay of migration and genetic drift. By incorporating temporal information, *i.e.*, the relative ages of alleles, Templeton *et al.* (1995) have shown that it is possible to discriminate ongoing (gene flow) from historical processes such as vicariant fragmentation.

In an effort to compare both approaches side by side, we studied intraspecific patterns of genetic variation in the freshwater mussel, *L. hydiana*. Phylogeographic analysis using the method of Templeton *et al.* (1995) indicated that historical population fragmentation is the most important force shaping genetic diversity in this species. Two distinct fragmentation events were identified. The oldest event was the separation of Arkansas River (LcfH 16, 17) populations from the remainder of populations located farther southwest in the Ouachita and Saline river drainages. A second, more recent fragmentation event was inferred between the upper Saline River and the remainder of the populations in the lower Saline, Ouachita, and Neches rivers.

Nested cladistic analysis provided no insight into processes shaping local population differentiation for haplotypes contained in nesting clade 1-2 because an interior one-step haplotype was not identified for this clade. We suspect that SSCP screening may have failed to detect a number of single nucleotide changes (based on their absence from the data set). Consequently, we could not estimate interior-tip ($D_cI - D_cT$ and $D_nI - D_nT$) values for nesting clade 1-2 because the putative one-step interior haplotype connecting clade 1-2 to the remainder of the parsimony network did not contain geographic or haplotype frequency information. While

TABLE 5

A comparison of inferences made about processes affecting spatial genetic variation of *L. hydiana* between nested cladistic analysis (GEODIS) and analysis of molecular variance (AMOVA)

Inference	GEODIS[a]	AMOVA[b]	Methods concordant?
Fragmentation of Arkansas River *vs.* other *L. hydiana*	Yes; nine mutational steps separating haplotypes	Yes; fixed differences among haplotypes	Yes
Fragmentation of Upper Saline *vs.* Lower Saline plus Ouachita	Yes (see Table 4)	No; low among-group variance relative to within-group variance suggests reduced local gene flow	No
Isolation by distance	No, at higher clade levels (see Table 4); equivocal at lower clade levels (interior-tip comparison impossible)	No; slope of regression not significant, low r^2 value observed	Yes

Evidence supporting each inference is included.
[a] Templeton *et al.* 1995; Posada *et al.* 1999.
[b] Excoffier *et al.* 1992.

significant geographic structuring can be inferred without interior-tip comparisons (via homogeneity testing), the process that likely shaped geographic structuring within nesting clade 1-2 cannot (see appendix in Templeton *et al.* 1995). Thus, inferential power of nested cladistic analysis may depend on the ability to resolve one-step mutational changes with a high degree of certainty.

AMOVA identified significant genetic structure at all levels in both hierarchical designs examined. The presence of a fixed genetic difference between Arkansas River populations, and the observation of large among-group relative to within-group genetic variance, were consistent with the cessation of gene flow resulting from allopatric fragmentation. This interpretation was congruent with nested cladistic analysis. In the second hierarchical analysis that examined variance partitioned among river drainages, within-group variance was much higher than among-group variance. Moreover, regression of pairwise $\Phi/(1-\Phi)$ and river distance indicated that population differentiation did not follow an isolation-by-distance model. Taken together, the second hierarchical analysis and pairwise regression using AMOVA suggested reduced gene flow at the local population level under an island model.

AMOVA did not provide strong support for a second fragmentation event because relatively little genetic variation is attributable to differences among river drainages (Arkansas River excluded). However, if the hierarchical design is changed to reflect a contrast between Upper Saline *vs.* the remainder of *L. hydiana* localities (inferred from nested cladistic analysis), AMOVA results are consistent with (but do not unequivocally support) a fragmentation event (among-group variance $\Phi_{CT} = 0.470 >$ within-group variance $\Phi_{SC} = 0.196$). These results illustrate a key difference between AMOVA and nested cladistic analysis. In AMOVA, the hierarchy is identified *a priori*, whereas in nested cladistic analysis the hierarchical groups emerge from the data via an objective (parsimony) criterion. Holsinger and Mason-Gamer (1996) proposed a method based on Wright's F_{ST} that produces a rooted tree (with sampling localities as OTUs) that allows any hierarchical structure to be revealed in the analysis. However, it is important to note that for any approach based on analysis of variance (*e.g.*, Wright's *F*-statistics, AMOVA), large *F*- or Φ-values can only be interpreted as a reduction in current gene flow (or effective population size). Therefore, the only unequivocal evidence of cessation of gene flow (as in a fragmentation event) is the observation of fixed allelic differences across sampling localities.

While we are very enthusiastic about the application of the nested cladistic method to analysis of population structure, several caveats must be considered in its interpretation. Perhaps most critical is the assumption that interior (ancestral) haplotypes are more geographically widespread than tip (derived) haplotypes under the null hypothesis of spatial panmixia. It is conceivable that selection could yield haplotype geographic patterns that violate this assumption. The assumption of neutrality in genetic markers studied is shared by AMOVA. Another apparent difficulty is the lack of resolution of the nested cladistic method when interior haplotypes are not detected, a result that could occur either because sample sizes are small or because the screening method employed is insensitive to certain mutational changes. The nested cladistic method requires interior-tip distances to infer processes structuring clades within a given nesting clade (see Templeton *et al.* 1995). Under these circumstances, more traditional analysis of variance-based approaches may possess more power to infer processes that shape population structure when localities are near migration / drift equilibrium. However, our comparison indicated that the nested cladistic method provided objective identification of population fragmentation with statistical support that went undetected by the more traditional analysis of variance approach. Thus, a combination of nested cladistic and analysis of variance approaches may best describe spatial patterns of genetic variation: the former to identify population fragmentation events, and the latter to characterize gene flow within fragmented regions.

Molecular genetic work was conducted in the laboratory of D. Kuhn, and we are especially grateful to him for technical advice and assistance in all phases of the project. J. Bielawski and R. Broughton provided helpful discussions, and two anonymous reviewers provided comments that greatly improved the manuscript. A. R. Templeton provided computer programs for data analysis. M. Trexler drafted Figure 1. Mussel specimens from Texas were provided by B. Howells of the Texas Parks and Wildlife Department. *L. powelli* were sampled under U.S. Fish and Wildlife Service permit SA 92-18. This study was funded by U.S. Department of Agriculture Forest Service Cooperative Agreement 19-91-074 to J.T.

LITERATURE CITED

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb *et al.*, 1987 Intraspecific phylogeography: the mitochondrial-DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. **18:** 489–522.

Bernatchez, L., and C. C. Wilson, 1998 Comparative phylogeography of nearctic and palearctic fishes. Mol. Ecol. **7:** 431–452.

Burch, J. B., 1975 *Freshwater Unionacean Clams (Mollusca: Pelecypoda) of North America.* Malacological Publications, Hamburg, MI.

Devereux, J., P. Haeberli and O. Smithies, 1984 A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:** 387–395.

Excoffier, L., P. E. Smouse and J. M. Quattro, 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131:** 479–491.

Garrido-Ramos, M. A., D. T. Stewart, B. W. Sutherland and E. Zouros, 1998 The distribution of male-transmitted and female-transmitted mitochondrial DNA types in somatic tissues of blue mussels: implications for the operation of doubly uniparental inheritance of mitochondrial DNA. Genome **41:** 818–824.

Gorman, O. T., 1992 Evolutionary ecology and historical ecology: assembly, structure, and organization of stream fish communities, pp. 659–688 in *Systematics, Historical Ecology, and North American Freshwater Fishes*, edited by R. L. Mayden. Stanford University Press, Stanford, CA.

Harris, J. L., and M. E. Gordon, 1990 *Arkansas Mussels*. Arkansas Game and Fish Commission Publication. Little Rock, AR.

Hellberg, M. E., 1996 Dependance of gene flow on geographic distance in two solitary corals with different larval dispersal capabilities. Evolution **50:** 1167–1175.

Holsinger, K. E., and R. J. Mason-Gamer, 1996 Hierarchical analysis of nucleotide diversity in geographically structured populations. Genetics **142:** 629–639.

Hudson, R. R., 1989 How often are polymorphic restriction sites due to a single mutation? Theor. Popul. Biol. **36:** 23–33.

Johnson, R. I., 1980 Zoogeography of North American Unionacea (Mollusca:Bivalvia) north of the maximum Pleistocene glaciation. Bull. Mus. Comp. Zool. Harv. Univ. **149:** 77–189.

Liu, H.-P., J. B. Mitton and S.-K. Wu, 1996 Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the freshwater mussel, *Anodonta grandis grandis*. Evolution **50:** 952–957.

Lydeard, C., M. Mulvey and G. M. Davis, 1996 Molecular systematics and evolution of reproductive traits in North American freshwater unionacean mussels (Mollusca:Bivalvia) as inferred from 16S rRNA gene sequences. Philos. Trans. R. Soc. Lond. B **351:** 1593–1603.

Mayden, R. L., 1988 Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. Syst. Zool. **37:** 329–355.

Nei, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Neigel, J. E., 1997 A comparison of alternative strategies for estimating gene flow from genetic markers. Annu. Rev. Ecol. Syst. **28:** 105–128.

Pennak, R. W., 1978 *Fresh-Water Invertebrates of the United States*, Ed. 2. John Wiley & Sons, New York.

Posada, D., K. A. Crandell and A. R. Templeton, 1999 GEODIS version 2.0. available via world-wide-web at URL = http://bioag.byu.edu/zoology/crandall_lab/programs.htm

Ricklefs, R. E., 1987 Community diversity: the relative roles of local and regional processes. Science **235:** 167–171.

Roff, D. A., and P. Bentzen, 1989 The statistical analysis of mitochondrial DNA polymorphisms: χ^2 and the problem of small samples. Mol. Biol. Evol. **6:** 539–545.

Rousset, F., 1996 Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics **145:** 1219–1228.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Skibinski, D. O. F., C. Gallagher and C. M. Benyon, 1994 Mitochondrial DNA inheritance. Nature **368:** 817–818.

Swofford, D. L., 1995 *PAUP* 4.0 BETA for Macintosh Power PC*. Sinauer Associates, Sunderland, MA.

Templeton, A. R., 1998 Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. Mol. Ecol. **7:** 381–397.

Templeton, A. R., E. Routman and C. A. Phillips, 1995 Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics **140:** 767–782.

Tyler, B. M., H. Förster and M. D. Coffey, 1995 Inheritance of avirulence factors and restriction fragment length polymorphism markers in outcrosses of the oomycete *Phytophthora sojae*. Mol. Plant-Microbe Interact. **8:** 515–523.

van der Schalie, H., and A. van der Schalie, 1950 The mussels of the Mississippi River. Am. Midl. Nat. **44:** 448–466.

Zouros, E., A. O. Ball, C. Saavedra and K. R. Freeman, 1994 Mitochondrial DNA inheritance. Nature **368:** 818.

Communicating editor: M. K. Uyenoyama