Evolution, 48(6), 1994, 1799-1809

PARSIMONY, MOLECULAR EVOLUTION, AND BIOGEOGRAPHY: THE CASE OF THE NORTH AMERICAN GIANT SALAMANDER

ERIC ROUTMAN[1], ROSALIND Wu, AND ALAN R. TEMPLETON
Department of Biology, Washington University, St. Louis, Missouri 63130

Abstract. - To draw biogeographic conclusions about the Central Highlands region of the United States, we reconstructed the phylogeny of hellbender *(Cryptobranchus alleganiensis)* populations from restriction-site variation in mtDNA. We were unable to root the phylogeny using an outgroup and therefore could not weight restriction-site gains more heavily than site losses. As a result, maximum parsimony results in low phylogenetic resolution because of high levels of homoplasy in the data set. Use of a recently published algorithm based on an explicit model of molecular evolution yielded much greater resolution of the mtDNA relationships. This phylogeny indicates the two subspecies ofhellbenders are paraphyletic with respect to one another. Hellbenders found in the southern Ozarks (C. *a. bishopi)* are either most closely related to populations of C. *a. alleganiensis* inhabiting the Tennessee River drainage or are so divergent that phylogenetic affinities are undetectable. Extremely low levels of divergence among mtDNA haplotypes found in populations from Pennsylvania, Indiana, Illinois, and the northern Missouri Ozarks suggest a recent, probably post-Pleistocene, invasion of this region from a refugium in one of these areas. Biogeographic hypotheses of the causes and timing of hellbender distributions differ significantly from those postulated from analyses of fish species relationships. Possible reasons for the discrepancy are discussed.

Key words. -Biogeography, *Cryptobranchus,* molecular evolution, phylogeny.

Received January 3, 1994. Accepted May 31, 1994.

Explaining the distributions of organisms is one of the principal goals of both ecology and evolutionary biology. Evolutionary approaches focus on the contribution of historical events as a force determining distributions, using the principles of phylogeny reconstruction and biogeography (for a review, see Humphries and Parenti 1986). The application of these principles to the study of intraspecific populations was pioneered by Avise and colleagues (Bermingham and Avise 1986; Avise et al. 1987). Intraspecific biogeography is an example of the increasing awareness that historical information is not only valuable for understanding macroevolution but may shed light on microevolutionary processes as well. Conversely, the use of explicit population-genetic models of the evolutionary process can be useful for phylogenetic reconstruction.

Intraspecific biogeography involves the joint consideration of the phylogenetic and geographic distributions of populations within a species. DNA sequencing or r striction mapping are usually more informative than morphological data sets, because for most species these populations are very closely related. This field has rescued widespread species from the fate of being biogeographically "uninformative" (Nelson and Platnick 1978). Many widespread species consist of geographically distinct populations that also differ genetically for at least some molecular markers (Avise et al. 1987). These subpopulations can then be treated the same as endemic species in more traditional biogeographic studies, either by comparing phylogenies of several widespread species to find areas of concordance (Bermingham and Avise 1986; Avise 1992), or by comparing population relationships to those predicted by geologic hypotheses.

However, intraspecific molecular phylogenies suffer from problems that often make the use of traditional methods of phylogeny reconstruction difficult. One problem is branch lengths between nodes of the phylogeny are expected to be short. Methods like maximum parsimony perform best when internodal branch lengths are long relative to terminal branches (Felsenstein 1978). Secondly, independent rooting of the phylogeny is often not possible. Rooting is important because it allows the identification of natural groups, and is necessary to differentially weight certain types of character-state changes (such as restriction-site gains versus losses) that can alter the topology of the phylogeny. Because populations are often closely related, rapidly evolving characters,

[1] Present address: Department of Biology, San Francisco State University, 1600 Holloway Avenue, San Francisco, California 94132-2401.

such as mtDNA or the intergenic spacer ofrDNA, are used to estimate the phylogeny. However, rapid evolution, combined with a limited number of character states, can obviate the use of an outgroup to root the tree, because the molecule being studied is rapidly changing in the outgroup as well. This can result in autapomorphies of the outgroup appearing as synapomorphies in the ingroup.

The second problem is particularly acute with restriction-site data. Restriction sites have only two character states-present and absent. In addition, in closely related organisms, sites are considerably more likely to be lost than gained (Templeton 1983). This results in the expectation that, for sites "present" in at least some members of the ingroup, the outgroup will accumulate "absent" character states due to independent loss. This can result in an incorrect or ambiguous rooting of the phylogeny. In cases where the phylogeny-reconstruction algorithm returns a message that the phylogeny cannot be rooted with the outgroup, investigators commonly include the outgroup as part of the ingroup and root the phylogeny with some other criteria (e.g., Echelle and Dowling 1992). The above discussion suggests this may not always be appropriate. A test for outgroup suitability in the case of restriction-site data will be presented elsewhere.

To overcome these obstacles, other criteria are necessary to help decide group membership, such as using explicit population-genetic models of molecular evolution to justify decisions about the use of parsimony to estimate the phylogeny. Here we use a method based on an explicit model of restriction-site evolution to improve our biogeographic study of the giant salamander, *Cryptobranchus alleganiensis* (also known as the hellbender). The goals are to propose hypotheses about the phylogenetic relationship of hellbender populations in the Central Highlands of the United States and to compare these hypotheses to those derived from other taxa found in the same areas. These hypotheses will provide the basis for future biogeographic studies of the region.

The Central Highlands

The Central Highlands of the eastern United States is a most interesting region for intracontinental biogeographic investigation because of its biological diversity. It is divided into the Eastern Highlands, consisting of the Appalachian Mountains east of the Mississippi River, and the Interior Highlands, including the Ozark Plateau and Ouachita Mountains west of the Mississippi. These two major upland regions are characterized by mountainous topography and ecology and are separated from each other and from the rest of the continent by lowlands with very different ecological conditions. The highlands, characterized by relatively small streams with clear water and high flow gradients, are important to the distributions of aquatic organisms. Intervening lowland rivers tend to be larger and slower, with high siltation. For many upland aquatic organisms, lowland rivers represent considerable barriers to dispersal. Current distributions of these organisms probably reflect past geologic events rather than present-day dispersal patterns.

The most important conclusion of extensive reviews of the geologic history of Central Highlands rivers (Ray 1974; Wiley and Mayden 1985; Robison 1986; Mayden 1987, 1988) is that the present physiognomy of the region is due to geologic activity that occurred just before and during the Pleistocene. Rivers of the region were dramatically affected by Pleistocene glaciation events. The direction of flow of many contemporary rivers, e.g., the Allegheny, was reversed, and other drainages were created or destroyed. Flow rate and turbidity were also affected. The relationships among current populations of aquatic organisms may reflect pre-Pleistocene affinities or postglaciation phenomena. Studies of the region's fishes suggest the answer is a mixture ofboth (Wiley and Mayden 1985; Mayden 1987, 1988).

Several biogeographic investigations of Central Highlands aquatic organisms have focused on shared community structure (Pflieger 1971) or interspecific vicariance biogeography (Wiley and Mayden 1985; Mayden 1988). Surprisingly, few intraspecific biogeographic studies have been conducted (Wiley and Mayden 1985; Rogers and Cashner 1987; Grady et al. 1990). Several widespread species in a variety oftaxa, including crayfish, mollusks, fish, amphibians, and plants, are restricted to rivers in the Central Highlands. Comparing the phylogenies of these species would help elucidate the effects of past geologic events on the current biological makeup of the region.

Biology of Cryptobranchus alleganiensis

Cryptobranchus alleganiensis is an obligately aquatic salamander found in clear, fast-flowing rivers in the eastern United States, from southern New York to northern Georgia, Alabama, and

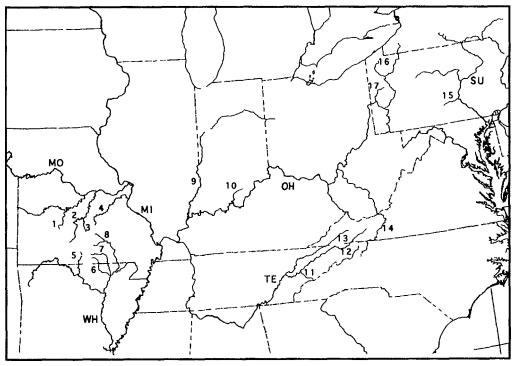


FIG. 1. Map of rivers where *Cryptobranchus alleganiensis* populations were sampled: (!) Niangua River, (2) Gasconade River, (3) Big Piney River, (4) Meramec River, (5) North Fork of the White River, (6) Spring River, (7) Eleven Point River, (8) Current River, (9) Wabash River, (10) Blue River, (11) Little River, (12) Beaverdam Creek, (13) Copper Creek, (14) New River, (15) Sherman Creek, (16) French Creek, (17) Slippery Rock Creek. Major drainages are: MO, Missouri River; Ml, Mississippi River; WH, White River; OH, Ohio River; TE, Tennessee River; SU, Susquehanna River.

Mississippi, and from the western parts of West Virginia, Virginia, and the Carolinas to central and southern Missouri (Nickerson and Mays 1973). The main body of the range includes tributaries of the Ohio, Cumberland, and Tennessee drainages of the Appalachian Mountains, but disjunct populations are found in the Ozark Plateau and the Susquehanna drainage. The hellbenders that inhabit rivers flowing south out of the Ozarks are considered a separate subspecies, C. *a. bishopi* (Grobman 1943), whereas the northern Ozark populations and all other populations are thought to belong to the nominate subspecies, C. *a. alleganiensis.* The Susquehanna drains into Chesapeake Bay, and its populations are therefore isolated from the other populations, which inhabit rivers draining ultimately into the Mississippi River.

Hellbenders require large areas of submerged rock as diurnal refugia and nesting sites, and are not found in water with high silt loads. Therefore, they can only disperse between upland areas when lowland rivers are relatively clear. Furthermore, because hellbenders fertilize eggs externally, single gravid females are unable to found new populations unless they comigrate with at least one male. Thus, new populations are unlikely to result from extremely rare dispersal by individuals through unsuitable habitat, but rather are probably founded when the lowland rivers have much reduced silt loads. This specialized ecology means the distribution and relationships of C. *alleganiensis* populations should be indicative of past climatic and geologic events.

MATERIALS AND METHODS

Two hundred fifty-two hellbenders were collected from various rivers throughout the species' range. The collection sites are shown in figure I. Table 1 lists sample sizes and drainage relationships. Blood from an incision in the tail was dripped into cryotubes containing STE buffer (Maniatis et al. 1982) and frozen in liquid nitrogen. Procedures for extraction, restriction, electrophoresis, and visualization of DNA are described in Routman (1993). A preliminary

TABLE I. Collecting localities, state where collection was made, sample sizes, and drainages. Drainage pattern refers to the sequential listing of tributaries, starting with the river where the collection was made and ending with the last river before the confluence with the ocean. State abbreviations are: AR, Arkansas; IL, Illinois; IN, Indiana; MO, Missouri; PA, Pennsylvania; TN, Tennessee; VA, Virginia.

Big Piney River (MO)	65	Gasconade-Missouri-Mississippi
Gasconade River (MO)	13	Missouri-Mississippi
Niangua River (MO)	26	Osage-Missouri-Mississippi
Meramec River (MO)	42	Mississippi
Current River (MO)	10	Black-White-Mississippi
Eleven Point River (MO)	I	Black-White-Mississippi
Spring River (AR)	7	Black-White-Mississippi
North Fork of the White River (MO)	17	White-Mississippi
Little River (TN)	12	Tennessee-Mississippi
Beaverdam Creek (TN)	8	Holston-Tennessee-Mississippi
Copper Creek (VA)	16	Clinch-Tennessee-Mississippi
New River (VA)	2	Kanawha-Ohio-Mississippi
French Creek (PA)	13	Allegbeny--Ohio-Mississippi
Slippery Rock Creek (PA)	6	Beaver-Ohio-Mississippi
Blue River (IN)	6	Ohio-Mississippi
Wabash River (IL)	I	Ohio-Mississippi
Sherman Creek (PA)	7	Susquehanna

subset of these data appeared in Templeton et al. (1990) and in Routman (1993).

Each mtDNA molecule that possesses a unique combination ofrestriction sites is considered an allele, or haplotype. In the phylogenetic analyses, haplotypes were considered operational taxonomic units (OTUs), and restriction sites were characters coded as present or absent. Parsimony analysis was conducted using PAUP (Swofford 1991). Twenty heuristic analyses, each with 10 separate random-addition taxa inputs, were performed. Characters and character states within characters were weighted equally. Maxtrees was set at 100 and allowed to increment by 100 as necessary.

The unique insights of an explicit model of evolution are introduced into the analysis by a method designed specifically for intraspecific restriction-site data (Templeton et al. 1992). The method combines maximum parsimony with Bayesian probability analyses to calculate the statistical limits of parsimony for a given data set. Full details of the algorithm are given in Templeton et al. (1992), so we shall only summarize them here. The basic goal is to evaluate the probability of multiple mutations within restriction sites. If undetected multiple mutations are not likely to have occurred within restriction sites, the use of parsimony to establish relationships among the haplotypes is justified (Templeton et al. 1992). The maximum number of restriction-site differences that can separate haplotypes before the probability of undetected multiple mutation events becomes too large (see below) is found with an algorithm inspired by the theoretical work of Hudson (1989), who calculated the probability of two haplotypes chosen at random having undetected restriction-site differences. The Templeton et al. method takes into account the overall restriction-site similarity of the specific pair of haplotypes being compared, which did not concern Hudson. Because the Templeton et al. procedure is conditional upon a specific pair of haplotypes, it makes no assumptions about the population structure and genetic equilibrium of the sampled population. The probability of parsimony *(P,* the probability ofno undetected mutations in the restriction sites sampled) is calculated for haplotypes separated by a single restriction site, then for those separated by two restriction sites, and so on until *P* becomes too small (e.g., <0.95). Parsimony is then used to reconstruct relationships within those subsets of haplotypes for which the probability ofundetected multiple mutations is less than 0.05. In contrast to the global parsimony used by PAUP, the Templeton et al. algorithm uses parsimony in a more localized manner. Local parsimony is more appropriate for data sets with relatively high rates of change among a limited number of character states. In these situations, homoplasy is expected for the data set overall but should be much less common among closely related taxa.

RESULTS

Table 2 gives the restriction-site profiles of the mtDNA haplotypes. This data set is essentially the same as that given in Routman (1993), with the addition of haplotypes P2 and I1 resulting from the sampling of populations from the Blue River, Wabash River, and French Creek, additional sampling of Slippery Rock Creek, and the addition of a sample from the Eleven Point River. A map of the mtDNA showing the relative locations of the restriction sites can be found in Routman (1993). A total of 17 different haplotypes was found. Population haplotype frequencies are found in the Appendix.

No haplotype was widespread among the populations sampled. Indeed, only a few haplotypes were found in more than one population, and $F_{,,}$ values for hellbenders are among the highest ever recorded (0.865 for the entire range, 0.779 for Missouri River tributaries-Routman 1993). Mitochondrial DNA variance in these populations is almost entirely explained by the between-population component, and because restricting inference to geographically adjacent populations has no significant effect on $F_{,,,}$ hellbender populations in different rivers probably rarely, if ever, exchange migrants (Routman 1993). This conclusion is consistent with results of mark-recapture studies showing low within-river movement and considerable philopatry in this species (Nickerson and Mays 1973; Peterson 1987). This geographic isolation means that the mtDNA haplotypes probably reflect the events that resulted in the founding of the populations in which they occur. The restriction-site differences among the haplotypes can 'be overlaid on a map to simultaneously display the mutational and geographic relationships of the mtDNA genomes of these hellbender populations (fig. 2).

Several notable aspects of the mutational relationships stand out. Some geographically close populations are separated by a large number of restriction-site differences. For example, at least

TABLE 2. Polymorphic restriction sites for hellbender mtDNA haplotypes. Haplotype designations consist of the first initial of the state where the collection was made (**M,** Missouri; P, Pennsylvania; T, Tennessee; V, Virginia) followed by a separate number for each haplotype. Only variable sites are listed (1 = present, 0 = absent). Restriction enzyme designations are as follows: B = *BamHI,* C = *BclI,* G = *BglII,* Bs = *BstEII,* E = *EcoRI,* Ev= *EcoRV,* N = *Neal,* P = *Pstl,* P = *Pvull,* S = *Sacl,* T = *Stul,* X = *Xbal,* **M** = *Xmnl.*

Hap-lo-type	Restriction sites												
	B	C	G	Bs	E	Ev	N	P	V	**S**	T	**X**	**M**
	I 2	2 3 4 5 6		2		2 3 4 5 6 7 8	I 2 3		2	I 2	2		I 2
Ml	1 0	1 0 0	0	1 0	0	1 0 0 0 0 0 0	1 0	0	1 1	1 1	1 1	0	0
M2	1 1	1 0 0	0	1 0	0	1 0 0 0 0 0 0	1 0	0	1 1	1 1	1 1	0	0
M3	1 1	1 0 0	0	1 1	0	1 0 0 0 1 0 1	1 0	0	1 1	1 1	1 1	1	0
M5	1 1	1 0 0	0	1 0	0	1 0 0 0 1 0 1	1 0	0	1 1	1 1	1 1	1	0
M6	1 0	1 0 0 1	0	1 0	0	1 0 0 0 1 0 1	1 0 1	0	1 1	1 1	1 1	1	0
M7	1 0	0 1 1 0	0	0 0	0	0 0 0 1 0 1 0	0 1 0	0	1 0	0 0	0 1	1	0
MS	1 0	1 0 1 1 0	1	0 0	0	0 0 0 1 0 1 0	1 1 0	0	1 0	0 0	0 1	1	1 1
M9	0 1	0 0 0 1 1	0	0 0	0	0 0 1 0 1 0 0	0 1 0	0	0 1	0 1	0 0	1	1 0
Pl	1 0	1 1 0 0 1	0	1 0	1	1 1 1 0 1 0 1	1 0 1	0	1 1	1 1	1 1	1	0 1
P2	1 1	1 1 0 0 1	0	1 0	0	1 1 1 0 1 0 1	1 0 1	0	1 1	1 1	1 1	1	0 1
11	1 0	1 1 0 0 1	0	1 0	0	1 1 1 0 1 0 1	1 0 1	0	1 1	1 1	1 1	1	0 1
Tl	1 0	0 0 0 1 0	0	0 0	0	0 0 0 1 0 1 0	1 0 1	0	1 1	0 0	0 1	1	0 0
T2	1 0	0 0 0 1 0	0	0 0	0	0 0 0 1 0 0 0	1 0 1	0	1 1	0 0	0 1	1	0 0
T3	1 0	0 0 0 1 0	0	0 0	0	0 0 0 1 0 1 0	1 0 1	0	1 1	0 0	0 0	1	0 0
T4	1 0	0 0 0 1 0	0	0 0	0	0 0 0 1 0 0 0 1	1 0 1	0	1 1	0 0	0 0	1	0 1
VI	1 0	1 0 0 1 0 1	0	0 0	0	0 0 0 1 0 0 0 0	1 0 0	1	1 0	1 1	0 1	1	0 0
V2	0 0	1OO11O	O	OO	O	OOO 1 0 1 0 1	1 0 1	0	1 0	1 1	0 1	1	0 0

15 restriction-site differences separate mtDNA of hellbenders in the Current River and Eleven Point River from those in the Spring River and the North Fork of the White River in the southern Ozarks. On the other hand, only 2 to 7 restriction-site differences separate haplotypes found in hellbenders from Pennsylvania, Indiana, and Illinois from those found in the northern Ozarks. The number of restriction-site differences and the percentage of nucleotide dissimilarity for each pairwise comparison of haplotypes are given in table 3. Sequence divergence among haplotypes ranges from 0.2% to 4.7%.

Three hundred fifty-three most parsimonious trees (length = 53, consistency index = 0.673) were resolved with the PAUP algorithm. The strict-consensus tree is shown in figure 3. Because the mtDNA of the nearest relatives of *Cryptobranchus,* the giant salamanders of Asia (genus *Andrias),* is too divergent to serve as an outgroup (Routman unpubl. data), the tree is unrooted. The only groups that can be unambiguously supported are ([Ml,M2],[{M3,M5,M6},{Pl,P2,Il }]), and (M7,M8). This conclusion assumes the root does not lie within any of these groups. Because of the large genetic distances between these sets of haplotypes, we argue that they are probably monophyletic groups, that is, the true root is most likely to be found on one of the long branch-

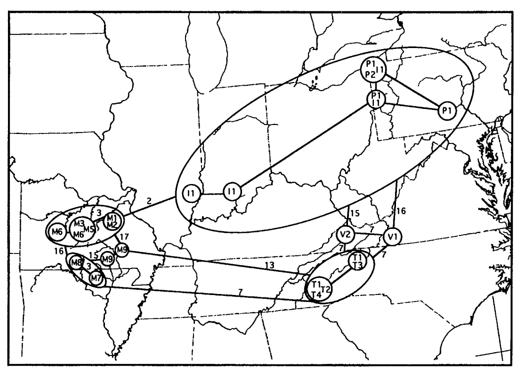


FIG. 2. Restriction-site differences among hellbender mtDNA haplotypes overlaid on a map of the eastern United States. Numbers next to lines connecting haplotypes or haplotype groups indicate the minimum number of restriction-site differences between haplotypes (absence of a number indicates shared haplotypes). Localities can be identified by their haplotype profiles as follows: French Creek-Pl,P2,Il; Slippery Rock Creek-Pl,11; Sherman Creek-Pl; Blue River-II (east); Wabash River-II (west); Meramec River-Ml,M2; Big Piney/Gasconade rivers-M3,M5,M6; Niangua River-M6; North Fork of the White River-M8; Spring River-M7; Eleven Point River-M9 (west); Current River-M9 (east); Little River-Tl,T2,T4; Beaverdam Creek-Tl,T3; Copper Creek-V2; New River-VI.

TABLE 3. Percentage of sequence difference (upper matrix) and number of restriction-site differences (lower diagonal) for each pairwise comparison of *Cryptobranchus alleganiensis* mtDNA haplotypes.

mtDNA	M1	M2	M3	M5	M6	II	P1	P2	M7	MS	M9	Tl	T2	T3	T4	VI	V2
Ml		0.2	1.0	0.8	0.6	1.0	1.2	1.2	3.8	3.9	4.7	3.1	2.9	3.4	2.9	3.1	2.8
M2	I		0.8	0.6	0.8	1.2	1.4	1.0	4.0	4.1	4.4	3.3	3.1	3.6	3.1	3.3	3.0
M3	5	4		0.2	0.4	0.7	0.9	0.5	4.3	4.4	4.1	3.6	3.4	3.9	3.4	3.6	3.3
MS	4	3	I		0.2	0.6	0.7	0.4	4.1	4.2	3.9	3.4	3.2	3.7	3.2	3.4	3.1
M6	3	4	2	I		0.4	0.6	0.6	3.9	4.0	4.2	3.2	3.0	3.5	3.0	3.2	2.9
II	5	6	4	3	2		0.2	0.2	4.3	4.4	4.1	3.6	3.4	3.9	3.4	3.6	3.3
Pl	6	7	5	4	3	I		0.4	4.5	4.6	4.3	3.8	3.6	4.1	3.6	3.8	3.5
P2	6	5	3	2	3	I	2		4.5	4.6	3.8	3.8	3.6	4.1	3.6	3.8	3.5
M7	16	17	19	18	17	19	20	20		0.7	3.8	1.7	2.0	2.0	2.0	2.2	2.4
MS	17	18	20	19	18	20	21	21	3		3.9	1.9	2.1	2.1	2.1	2.4	2.6
M9	19	18	18	17	18	18	19	17	15	16		3.6	3.4	3.4	3.4	4.2	3.8
Tl	13	14	16	15	14	16	17	17	7	8	14		0.2	0.2	0.7	2.0	1.7
T2	12	13	15	14	13	15	16	16	8	9	13	I		0.5	0.5	1.7	2.0
T3	14	15	17	16	15	17	18	18	8	9	13	I	2		0.5	2.3	2.0
T4	12	13	15	14	13	15	16	16	8	9	13	3	2	2		2.3	2.5
VI	13	14	16	15	14	16	17	17	9	10	16	8	7	9	9		1.7
V2	12	13	15	14	13	15	16	16	10	11	15	7	8	8	10	7	

es of the tree. A tree containing these clades is significantly more parsimonious than midpoint-rooted trees in which these clades are broken up ($P < 0.05$, Wilcoxon signed-rank test; Templeton 1983). Note that homoplasy is sufficiently great so that even Tennessee River haplotypes Tl, T2, T3, and T4 (separated by a maximum of three restriction sites) do not form a monophyletic group. Our inability to root this tree means that the homoplasy cannot be reduced by weighting site gains more heavily than site losses.

Use of the Templeton et al. algorithm allows further resolution of the haplotype relationships.

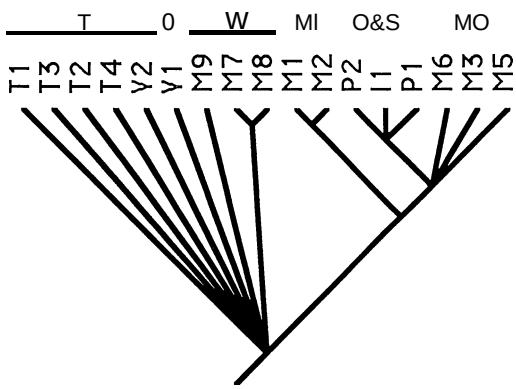
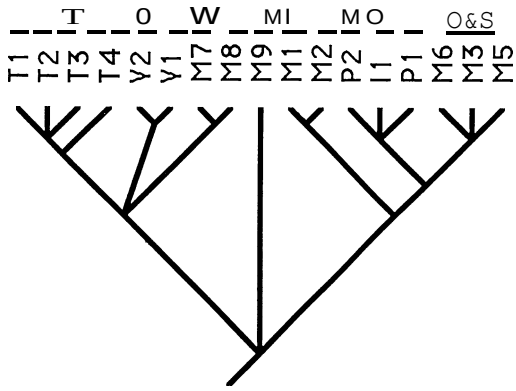


FIG. 3. Maximum parsimony phylogeny of hellbender mtDNA haplotypes. Letters above haplotypes represent major drainages where the haplotypes are found. Information about smaller drainages can be found in figures I and 2, table I, and in the text. T, Tennessee River; 0, Ohio River; **W,** White River; **MI,** Mississippi River; MO, Missouri River; S, Susquehanna River.

The first step in the application of the algorithm is to calculate H, the probability that two haplotypes picked at random from the sample differ by at least one undetected mutation in the set of restriction sites scored (Hudson 1989). For our data set, $H = 0.219$. This high probability indicates that maximum parsimony is inappropriate for the data set as a whole. However, it is still likely that parsimony can be used for subsets of the data. We estimate the limits of parsimony by applying equation (8) of Templeton et al. (1992) to haplotype pairs that differ by j restriction sites and share m restriction sites. (A parameter to account for transition/transversion bias is included in the model. We set this parameter so transitions are considered far more common than transversions, as is expected for closely related taxa. This makes the analysis more conservative, in that the ability to determine that no undetected mutations are present in the sample is decreased, making it more difficult to determine relationships among haplotypes.) The parameter j can range from 1 to the maximum number of restriction-site differences between haplotypes. The parameter m can range from zero to $N - j$, where N is the total number of restriction sites sampled. The most efficient estimation procedure is to begin with the haplotype pair with the lowest values of j and m. If the probability of no undetected mutations, P, is greater than or equal to 0.95, then parsimony is appropriate for all haplotype pairs that differ by this value of j. If $P < 0.95$, we proceed to haplotype pairs with successively higher values of m

for the same j until $P \geq 0.95$ or we run out of contrasts for that value of j. Clades are formed by uniting all haplotypes for which $P \geq 0.95$. We then proceed to the next higher value of j and repeat the process. Unresolved polytomies arise for two reasons: first, when $P < 0.95$ at the level of connection between haplotype groups; second, when $P \geq 0.95$, but multiple connections are possible.

In the hellbender mtDNA data set, the following haplotype pairs are separated by a single restriction site difference: Ml-M2, M3-M5, M5-M6, Pl-11, P2-Il, Tl-T2, andTl-T3. Ofthese "one-step" pairs, two pairs, (Tl-T2 andTl-T3), share the fewest number of restriction sites (35 sites), yielding a value of $P = 0.987$. Because the rest of the one-step pairs share more than 35 sites, all one-step relationships are more parsimonious than those requiring two or more steps. Similarly, all two-step relationships (see table 3) are more parsimonious than those of greater than two steps ($P \geq 0.960$). At this point, the algorithm allows formation of the following clades: (Ml,M2), ([M3,M5,M6], [Il,Pl,P2]), and ([Tl,T2,T3],T4). When the procedure is repeated for three-step haplotype pairs, $P < 0.950$ for some pairs (M1-M6, M2-M6, M5-M2, and M7-M8). This indicates that when haplotypes differ by as few as three restriction sites, undetected mutations are likely. However, the Templeton et al. algorithm also allows estimation of the probability ofno more than one undetected mutation (step 4 of Templeton et al. 1992). For each of the above three-step haplotype pairs, this probability is greater than 0.95. Therefore, three-step or four-step relationships are parsimonious for these haplotypes, and connections involving five or more steps are significantly less likely. This result allows us to (1) unite the (Ml,M2) clade with the ([M3,M5,M6], [11,Pl,P2]) clade, because the next closest connection involves 12 restriction-site differences, and (2) form the clade (M7,M8). The next value of j to be tested is 7. For $j = 7$, P 0.688. However, the probability of no more than one undetected mutation is 0.954. Because this means that seven-step or eight-step relationships are significantly more likely than those involving nine or more steps, the following clade can be formed: ([M7,M8], ([Tl,T2,T3],T4), [Vl,V2]). The only relationships left to be determined are those among (Ml,M2), ([M3,M5,M6], [Il,Pl,P2]), haplotype M9, and ([M7,M8], [Tl,T2,T3,T4], [Vl,V2]). Table 3 shows that members of these three clades differ from each other by a minimum of 12 restriction sites. P is well under 0.95 for any of these 12-step connections, as is the probability of only one additional undetected mutation. Therefore, the relationships among these three clades are left unresolved. Figure 4 shows the cladogram that results from the use of the algorithm.



Fm. 4. Phylogeny of hellbender mtDNA haplotypes obtained using the Templeton et al. (1992) algorithm. River designations are as in figure 3.

DISCUSSION

The phylogenetic analysis of hellbender mtDNA haplotypes reveals some surprising relationships. One of the most notable is the low genetic divergence among *Cryptobranchus alleganiensis alleganiensis* populations from the rivers (French Creek, Slippery Rock Creek, Blue River, Wabash River) that flow south into the Ohio River, and the populations from rivers (Meramec, Big Piney, Gasconade, Niangua) flowing north out of the Ozark Mountains into the Missouri or Mississippi rivers. The nucleotide difference among the haplotypes found in this region ranges from 0.4% to 1.4%, whereas the minimum difference between these haplotypes and haplotypes found in other populations of C. *a. alleganiensis* is 2.8%. This close relationship suggests that the invasion of the Ozarks from the Ohio River drainage or vice versa was relatively recent, compared to events that separated the ancestor of these populations from the others. One plausible explanation is that Pleistocene glaciation, which created the current Ohio River, wiped out hellbender populations in the rivers that existed north of the Ohio. As the glaciers receded, the rivers created were colonized

from northern Ozark populations that survived glaciation because the rivers they inhabited flowed toward the ice sheet and were unaffected (in their headwaters) by meltwater from the glacier. In their study of *Fundulus catenatus,* Grady et al. (1990) also found a close relationship between an Ohio River drainage population and those from the Ozarks (although, unlike in our study, the southern Ozark populations were included in this grouping).

Although the New River is (via the Kanawha River) a tributary of the Ohio River, the two hellbenders from our sample have a mtDNA haplotype (Vl) that is most closely related to mtDNA from populations from the Tennessee River drainage (haplotypes V2 and Tl-T4), and from the southern Ozarks (haplotypes M7 and M8). Like the northern Ozark rivers mentioned above, the New River is also a north-flowing river and may have served as a separate refugium. As a result, New River hellbender mtDNA is not closely related to that of the hellbenders occupying other Ohio River tributaries. The similarity of New River hellbender mtDNA to that of hellbenders from the Tennessee River may result from a pre-Pleistocene connection between these populations. The possibility of stream capture between these drainages is supported by both geologic and fish-distribution data (Ross and Carico 1963). Because of the small sample size (two individuals), and because the collection site is above Kanawha Falls, it is likely that the New River contains hellbenders that have mtDNA from the Ohio River drainage. Further sampling above and below Kanawha Falls should detect whether the New River contains a mixture of hellbender populations.

The sample of hellbenders from Sherman Creek, a Susquehanna River tributary, was monomorphic for one of the mtDNA haplotypes (Pl) found in hellbenders from other rivers in Pennsylvania. Recalling that the Susquehanna River drainage is presently totally isolated from the other rivers, the invasion of the Susquehanna was undoubtedly a very recent, post-Pleistocene event. This suggests that stream capture or a similar phenomenon is responsible. If so, other species which migrated between drainages at the same time should also show a similar pattern of low mtDNA divergence between populations from the Susquehanna drainage and those from the western Pennsylvania rivers.

The populations of C. *.a. bishopi* found in the rivers that flow south out of the Ozarks show some unexpected genetic affinities. Two extremely different mtDNA clades are found. One unites the hellbenders of the North Fork of the White River with those found in the Spring River. The other, consisting of only haplotype M9, unites the population sampled from the Current River with the single individual taken from the Eleven Point River. These clades differ by 16 to 18 restriction sites, or 3.8% to 3.9% nucleotide divergence. What makes this especially surprising is that the Spring River enters the Black River (of which the Current River is a tributary) within 5 km of the Eleven Point River. The Black River unites with the White River over 100 river kilometers downstream. The mtDNA ofhellbenders from the Spring River would be expected to be more closely related to that of hellbenders from the Current and Eleven Point rivers. The extreme divergence between the two clades suggests that the event separating the groups is very old. The final surprising result is, although one of the C. *a. bishopi* haplotypes (M9, from the Current and Eleven Point rivers) is so divergent that it cannot be confidently assigned a relationship to any of the other extant mtDNA haplotypes, the other C. *a. bishopi* haplotypes (M7, found in the Spring River, and M8, from the North Fork of the White River) are most closely related to the haplotypes ofhellbenders from the Tennessee River drainage. This relationship means that both subspecies of C. *alleganiensis* are paraphyletic.

The relationships among hellbender populations suggested by the mtDNA phylogeny invite comparison to phylogenies hypothesized for other species in the Central Highlands. Wiley and Mayden (1985) summarized much of the research on fishes of this region. They discussed 11 species groups with Central Highlands distributions that would allow at least partial comparison to C. *alleganiensis.* Although the fish and . hellbenders share areas of endemism, none of the species trees for fish is a good match for the *Cryptobranchus* population phylogeny. For example, for those species groups *(Notropus zonatus* and *Etheostoma variatum)* in which different taxa are found in the northern and southern Ozarks and in the Appalachians, the Ozark populations share a sister-group relationship. In the hellbenders, the northern Ozark populations are sister taxa to those in part of the Ohio River drainage, and a subset of the southern Ozark populations are most closely related to the Tennessee River populations.

Mayden (1988) conducted a biogeographic analysis of Mississippi River basin fish groups, based upon both the phylogenetic relationships within many fish taxa and the hypothesized distribution of ancestral taxa. The results of his analysis for the rivers from which hellbenders were sampled for this study are presented in figure 5A. Comparison to the phylogeny of the hellbender populations (fig. 5B) shows it would be difficult to find two more different topologies. Only the close relationship among the north-flowing Ozark rivers is found in both phylogenies. The lack of concordance between the two biogeographic analyses could be caused by a number of factors.

First, the distribution of hellbender mtDNA haplotypes may not be indicative of population relationships. In other words, the gene phylogeny is not a population phylogeny. If the mtDNA phylogeny reflects past geologic phenomena, then biogeographic conclusions drawn from the mtDNA tree are still sound. The fact that the phylogeny of other genetic markers may reflect the effects of earlier or more recent geologic events does not obviate the biogeographic conclusions drawn from the mtDNA phylogeny (although it means that the mtDNA phylogeny does not tell the "whole story" of the biogeography of the hellbenders). However, if the inconsistency of the mtDNA is due to sorting of mtDNA variation present in a widespread ancestral population, the mtDNA phylogeny is not a reflection of past geologic phenomena but is instead an artifact of random lineage extinction. This seems unlikely for two reasons. First is the extremely low gene flow mentioned earlier. Low gene flow means that geographic distribution of rapidly evolving molecules is likely to reflect relationships among populations and that widespread polymorphisms are unlikely (Neigel and Avise 1993). Widespread polymorphisms involving divergent haplotypes have never been documented in organisms with low gene flow (Avise et al. 1987). Second, there are large genetic distances between the major groups ofhaplotypes. Several authors have shown that the consistency probability between gene trees and population trees is expected to be high when internodal distances are large, especially if sample sizes are larger than three or four individuals per population (Pamila and Nei 1988; Takahata 1989; Hey 1991). In any event, the consistency between the organismal and mtDNA phylogenies for *C. alleganiensis* can be tested by comparing the mtDNA phylogeny to those estimated from a variety of nuclear genetic markers (Wu 1991).

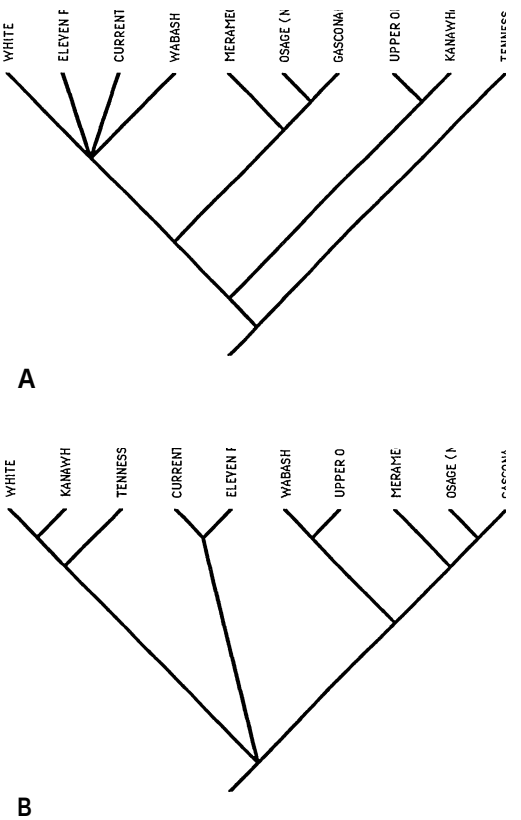


FIG. 5. A. Area cladogram of Central Highlands rivers based upon analysis of fish species relationships (Mayden 1988), including only drainages from which hellbenders were sampled for this study. B. Analogous area cladogram based on hellbender mtDNA phylogeny.

Second, the biogeographic conclusions of Mayden's (1988) study may not be correct. This could be due to violation of the assumptions of Mayden's method or to statistical error in the estimation of the phylogenies of the individual fish taxa (reconstructions were not tested statistically). The large number of fish taxa used in Mayden's study means that spurious correlations among phylogenies are an unlikely cause of error, however.

Third, the assumptions of the Templeton et al. algorithm may not be met by our data set. The major assumption is that a single mutation rate describes evolution in all the restriction sites in our data set. None of the restriction sites require numerous changes on any of the most par-

simonious trees, suggesting the mutation rates do not vary greatly. However, small rate variations are likely. The robustness of the algorithm to small deviations in mutation rate has not been explored.

Finally, it may be that the hellbender mtDNA phylogeny is correct but represents the results of dispersal, temporally different vicariance, or other unique events that have obscured the effects of the geologic history detectable in the fish study. We believe this is the most likely explanation. Differences in the ecology and dispersal ability between the hellbenders and the fish species studied by Mayden could result in the species responding differently to geologic phenomena. It will be interesting to see whether relationship patterns among other species with low vagility, an upland ecology, and dependence on rocky substrate are similar to those of the hellbenders.

Several of our conclusions (e.g., the relationships of southern Ozark and New River populations to those of the Tennessee River; the paraphyly of C. *a. bishopi)* can be drawn because the explicit model of Templeton et al. (1992) allows greater resolution of the phylogenetic tree in our case. We believe this approach will be useful for other biogeographic studies of widespread species, especially when out-group rooting is not feasible. Work on sculpins (C. Phillips pers. comm. 1994) and crayfish (K. Crandall pers. comm. 1994) is already in progress and may help elucidate the biogeographic history of the Central Highlands. The use of the Templeton et al. algorithm is certain to aid the analysis of these data sets, because the out-groups in these cases are relatively distantly related to the in-group.

ACKNOWLEDGMENTS

This work was supported by Missouri Department of Conservation Grant no. 62188 and National Institutes of Health (NIH) Genetics Training Grant no. 5-T32-GM08036, as well as NIH Grant no. R01-GM31571 (to A.T.). We thank many individuals, especially M. Hedin and C. Phillips for help in the field; K. Crandall, N. Georgiadis, and C. Phillips for comments on a preliminary version of the manuscript; and D. Foster, C. Iverson, T. Johnson, M. Nickerson, W. Redman, C. Shiffer, A. Sullivan, and **R. Wil**kinson for help with permits and locality data.

LITERATURE CITED

Avise,J.C. 1992. Molecularpopulationstructureand the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos 63:62-76.

Avise, J.C., J. Arnold, **R.M.** Ball, E. Bermingham, T. Lamb, J.E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18:489-522.

Bermingham, E., and J. C. Avise. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. Genetics 113:939-965.

Echelle, A. A., and T. E. Dowling. 1992. Mitochondrial DNA variation and evolution of the Death Valley pupfishes *(Cyprinodon,* Cyprinodontidae). Evolution 46:193-206.

Felsenstein. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27:401-410.

Grady, J. M., R. C. Cashner, and J. S. Rogers. 1990. Evolutionary and biogeographic relationships of *Fundulus catenatus* (Fundulidae). Copeia 1990:315-323.

Grohman, A. B. 1943. Notes on salamanders with the description of a new species of *Cryptobranchus.* Occasional Papers of the Museum of Zoology, University of Michigan 470:1-12.

Hey, J. 1991. The structure of genealogies and the distribution of fixed differences between DNA-sequence samples from natural populations. Genetics 128:831-840.

Hudson, R. R. 1989. How often are polymorphic restriction sites due to a single mutation? Theoretical Population Biology 36:23-33.

Humphries, C. J., and L. R. Parenti. 1986. Cladistic biogeography. Oxford University Press, New York.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mayden, R. L. 1987. Historical ecology and North American highland fishes: a research program in community ecology. Pp. 210-222 *in* W. J. Matthews and D. C. Heins, eds. Community and evolutionary ecology ofNorth American stream fishes. University of Oklahoma Press, Norman, Okla.

--. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. Systematic Zoology 37:329-355.

Neigel, J.E., and J. C. Avise. 1993. Application of a random walk model to geographic distributions of animal mitochondrial DNA variation. Genetics 135:1209-1220.

Nelson, G., and N. I. Platnick. 1978. The perils of plesiomorphy: widespread taxa, dispersal, and phenetic biogeography. Systematic Zoology 27:474-477.

Nickerson, M. A., and C. E. Mays. 1973. The hellbenders: North American "giant salamanders." Milwaukee Public Museum Press, Milwaukee, Wis.

Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5:568-583.

Peterson, C. L. 1987. Movement and catchability of the hellbender, *Cryptobranchus alleganiensis.* Journal of Herpetology 21:198-204.

Pflieger, W. L. 1971. A distributional study of **Mis**souri fishes. University of Kansas Museum ofNatural History Publication Series 20:226-570.

Ray, L. L. 1974. Geomorphology and Quaternary geology of the glaciated Ohio River Valley-aaa reconnaissance study. U.S. Geological Survey Professional Publications 826:1-7.

Robison, H. W. 1986. Zoogeographic implications of the Mississippi River basin. Pp. 267-285 *in* C. H. Hocutt and E. 0. Wiley, eds. The zoogeography of North American freshwater fishes. Wiley, New York.

Rogers, J. S., and R. C. Cashner. 1987. Genetic variation, divergence, and relationships in the subgenus *Xenisma* of the genus *Fundulus.* Pp. 251-257 *in* W. J. Matthews and D. C. Heins, eds. Community and evolutionary ecology of North American stream fishes. University of Oklahoma Press, Norman, Okla.

Ross, R. D., and J. E. Carico. 1963. Records and distribution offishes of the North, Middle, and South Forks of the Holston River, Virginia. Technical Bulletin 161 of the Virginia Agricultural Experiment Station, Blacksburg, Va.

Routman, E. 1993. Mitochondrial DNA variation in *Cryptobranchus alleganiensis,* a salamander with extremely low allozyme diversity. Copeia 1993:375-384.

Swofford, D. L. 1991. PAUP. Phylogenetic analysis using parsimony, Version 3.0. Illinois Natural **His**tory Survey, Champaign, Ill.

Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics 122:957-966.

Templeton, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. Evolution 37:221-244.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:6 I 9-633.

Templeton, A. R., K. Shaw, E. Routman, and S. **K.** Davis. 1990. The genetic consequences of habitat fragmentation. Annals of the Missouri Botanical Garden 77:13-27.

Wiley, E. 0., and R. L. Mayden. 1985. Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. Annals of the Missouri Botanical Garden 72:596-635.

Wu, C. I. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. Genetics 127:429-435.

Corresponding Editor: G. Vermeij

APPENDIX

Sample sizes and mtDNA haplotype frequencies for populations of *Cryptobranchus alleganiensis.*

Collection site	*N*	Haplotype frequencies		
Big Piney River	65	M5:0.877	M6:0.077	M3:0.046
Gasconade River	13	M5:l.000		
Niangua River	26	M6:1.000		
Meramec River	42	Ml:0.929	M2:0.077	
Current River	10	M9:1.000		
Eleven Point River	I	M9:1.000		
Spring River	7	M7:l.000		
North Fork of the White River	17	M8:1.000		
Little River	12	Tl:0.750	T2:0.166	T4:0.083
Beaverdam Creek	8	Tl:0.875	T3:0.125	
Copper Creek	16	V2:l.000		
New River	2	Vl:1.000		
French Creek	13	Pl:0.461	II :0.385	P2:0.154
Slippery Rock Creek	6	Il:0.833	Pl:0.167	
Blue River	6	Il:1.000		
Wabash River	I	Il:1000		
Sherman Creek	7	Pl:1.000		