Evolution, 57(9), 2003, pp. 2101–2118

LINEAR HABITATS AND THE NESTED CLADE ANALYSIS: AN EMPIRICAL EVALUATION OF GEOGRAPHIC VERSUS RIVER DISTANCES USING AN OZARK CRAYFISH (DECAPODA: CAMBARIDAE)

JAMES W. FETZNER, JR.[1,2] AND KEITH A. CRANDALL[1,3]
[1]*Department of Integrative Biology, Brigham Young University, 675 Widtsoe Building, Provo, Utah 84602*
[3]*Monte L. Bean Life Science Museum, Brigham Young University, Provo, Utah 84602*
E-mail: keith_crandall@byu.edu

Abstract.—The nested clade analysis can be extremely useful in testing for an association between genetic variation and geography and in explaining these observed patterns in terms of historical or contemporary population processes. The strength of this method lies in its ability to test a variety of processes simultaneously under a rigorous statistical framework. Indeed, many recent studies have used the nested analysis in a wide range of terrestrial and aquatic taxa. However, it has been suggested that riverine, riparian, or coastal species may be better examined using river (or coastal) distances rather than the standard geographic (great circle) distances among populations. It is thought that the standard geographic distances may not adequately describe the actual distances involved between populations of species inhabiting these one-dimensional (riverine) habitats. Therefore, we analyzed population data from an Ozark crayfish, *Orconectes luteus*, to examine the effects on the results of a nested clade analysis using river distances. In most cases, the haplotypes detected in this crayfish were unique to a particular drainage or a group of neighboring drainages, indicating very little movement of individuals among drainages. Five major population groups were detected, corresponding to many of the major river drainages sampled in this study. The two types of distance analyses obtain similar results for higher-level (older) clades, but differ in many of the inferences made for lower-level (younger) clades. However, we suggest that the comparison of both types of analyses for riverine species may enhance the process of elucidating historical and contemporary population processes, especially in cases where the transfer of individuals among different drainages are involved.

Key words.—Cambaridae, crayfish, genetic variation, linear habitats, nested clade analysis, phylogeography.

Received August 23, 2002. Accepted March 17, 2003.

The basic goal of population genetic studies is to describe how genetic variation is distributed within and among populations. However, additional information from gene genealogies has added a temporal perspective to phylogeographic analyses, which allows such studies to examine the geographic distribution of genetic variation in a historical framework. This combination of spatial and temporal information has enlightened many of our views of underlying population processes. Many phylogeographic methods fall short, however, because the detection of an association between haplotype distributions and geography does not necessarily lead to an explanation of this observed pattern (Templeton et al. 1995). One method developed to specifically address this problem is the nested clade analysis (NCA) of Templeton et al. (1987) and later extended by Templeton et al. (1995) and Templeton (1998, 2001).

The NCA is a powerful method for examining the geographic associations of haplotypes under a rigorous statistical framework. The method also goes further than other phylogeographic methods by explaining these associations in terms of contributions from either historical (e.g., fragmentation, colonization, or range expansion) or present-day (e.g., restricted gene flow) processes that have played a role in defining the currently observed patterns of population structure (Templeton et al. 1995; Templeton 1998).

Several other methods are also currently available that use temporal information contained in haplotype data. These include the analysis of molecular variance (AMOVA) of Excoffier et al. (1992) and the variance in ordered alleles method (N_{ST}) of Dumolin-Lapegue et al. (1997). However, the NCA allows for the analysis of more complex population processes than the N_{ST} method and provides a more objective assessment of geographic partitioning of haplotypes (and clades) than does the AMOVA method (Cruzan and Templeton 2000; Emerson et al. 2001; but see also Knowles and Maddison 2002).

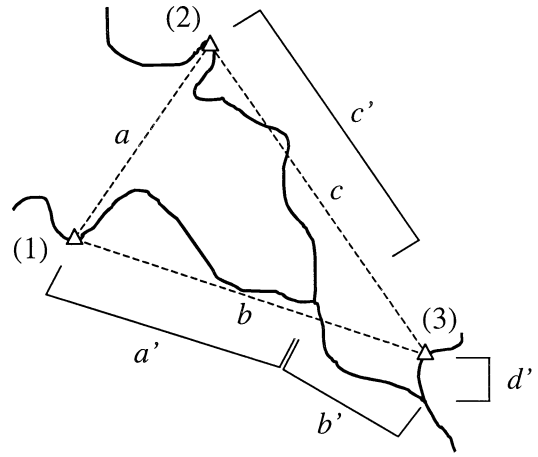
In addition, a few alternative methods have been used to infer population histories and include the mismatch pair distribution method of Rogers and Harpending (1992) and the skyline plot method of Pybus et al. (2000). However, both of these methods are limited in that they only deal with a single historical event or deal strictly with range expansions (or contractions).

While the NCA has been used in numerous analyses (e.g., Turner et al. 2000; Bernatchez 2001; Carbone and Kohn 2001; Sivasundar et al. 2001a), most such studies use geographic coordinates (latitude and longitude) to calculate the great circle distance between sampling sites or populations. For many riverine, riparian, or coastal species, however, these geographic distances may not adequately represent the actual distances separating such populations (Fig. 1). In such cases, user defined distances (hereafter referred to as ''river distances'') calculated by following river courses or coastlines may better reflect the actual distances that must be traversed if individuals were to migrate between populations.

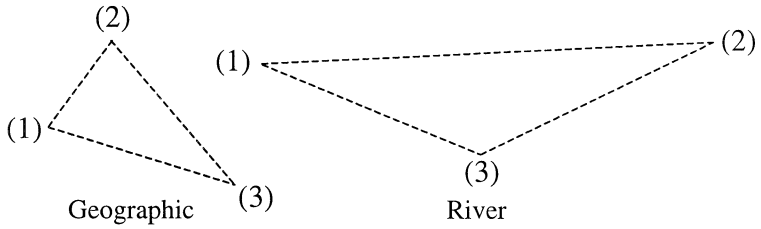
Linear river distances are often much larger than the geographic distances calculated among populations (see Fig. 2, Appendix 1). However, no empirical studies have been con-

[2] Present address: Section of Invertebrate Zoology, Carnegie Museum of Natural History, 4400 Forbes Avenue, Pittsburgh, Pennsylvania 15213; E-mail: fetznerj@carnegiemuseums.org.

A.



B.



C.

Population Distances	Geographic	River
(1) – (2)	*a*	*a'* + *c'*
(1) – (3)	*b*	*a'* + *b'* + *d'*
(2) – (3)	*c*	*c'* + *b'* + *d'*

FIG. 1. An illustration of the difference in estimating distances among populations of a riverine or riparian species using either geographic or river distances. (A) Demonstration of the different distances calculated among three hypothetical populations (△) from the same river drainage. (B) An illustration of the difference in distances calculated using the different methods. For geographic distances, the distance between populations 1 and 2 is the smallest, whereas for the river distances, populations 1 and 2 have the largest distance. (C) How these distances are calculated among the populations.

ducted to compare the effects these two different distances may have on results obtained from an NCA or whether the added work of calculating river distances is a necessary endeavor. Therefore, in this study, we compare the results of an NCA using both geographic and linear river distances to examine the population structure of the golden crayfish (*Orconectes luteus*).

The golden crayfish is one of the most common species of crayfish found in the Ozarks region of Missouri and lives in a variety of habitats including moderate- to fast-flowing streams, usually with rocky substrates. This species is currently distributed in four states (Fig. 3) including: northeastern Arkansas, east-central Kansas, southern and northeastern Missouri, and west-central Illinois (Wetzel and Poly 2000), although apparent misidentifications of older museum specimens coupled with new sample collections suggest its range may extend further north up the Mississippi River drainage than previously thought (J. E. Wetzel, W. E. Poly, and J. Fetzner, Jr., unpubl. data). This species occurs in most of the major river drainages in the Ozarks and provides a good test for the comparison of different population-to-population distance measures in a NCA.

Our objectives were: (1) to examine the population structure of the golden crayfish using mitochondrial 16S gene sequences; (2) investigate the processes contributing to the currently observed population structure of the golden crayfish through a NCA; and (3) to compare and contrast the results of two different types of NCA conducted on this riverine species, one using standard geographic distances (see below) and the other using linear river distances among populations.

MATERIALS AND METHODS

Population Samples

We analyzed a total of 393 individuals from 35 different populations of the golden crayfish (*O. luteus*) from throughout the species range in Missouri and Illinois, including most major river drainages where this species occurs (Fig. 3, see Appendix 2). Samples of abdominal or chela muscle were collected in the summers of 1998 and 2000 and frozen using

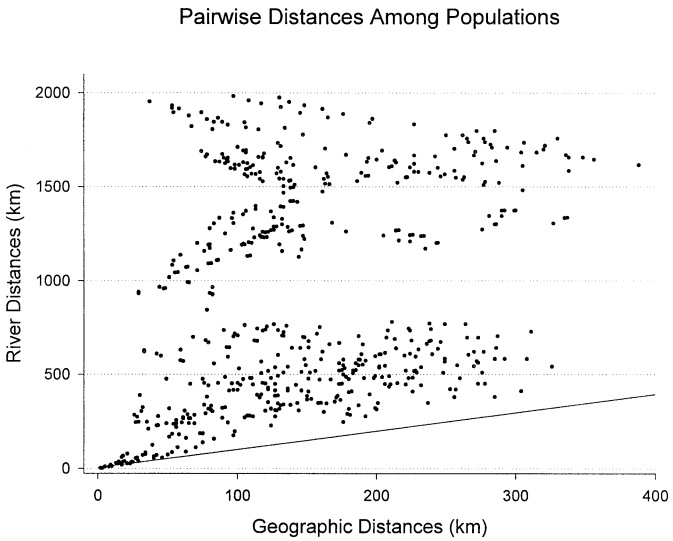


FIG. 2. Plot comparing all pairwise distances (km) among collection localities from this study for both geographic and river distances. The solid line indicates the expected distribution of points if the two distance measures examined were identical. The plot shows that in almost all cases the river distances far exceed the values estimated for the geographic distances among populations.

liquid nitrogen for transport back to the laboratory. After arriving at the laboratory, the samples were placed at −80°C until analysis. Voucher specimens from each population (one male and one female) were placed in the crustacean collection of the Monte L. Bean Life Science Museum located on the campus of Brigham Young University.

DNA Extraction, Amplification, and Sequencing

DNA was extracted using a high salt precipitation method previously described (Crandall et al. 1999). In brief, approximately 100 mg of tissue was placed into 900 μl of lysis buffer (10 mM Tris base, 100 mM EDTA, 2% SDS, pH 8.0) along with 9 μl of proteinase K (10 mg/ml). The samples were then allowed to incubate overnight at 58°C. The extraction continued after cooling the samples to room temperature and adding 4 μl of RNase A (20 mg/ml). The samples were then incubated at 37°C for 1 h, after which 300 μl of 7.5 M ammonium acetate was added and the samples were then vortexed and placed on ice for 15 min. The samples were then centrifuged at high speed for 5 min and the supernatant was then added to 900 μl of isopropanol. Samples were then inverted several times to precipitate the DNA and placed at −20°C overnight. The next day, the samples were centrifuged and the pellet was washed with 500 μl of 70% ethanol. After drying the pellet completely, 50 to 200 μl of TLE buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0) was added to resuspend the DNA. The quantity of DNA was then checked using a spectrophotometer and dilutions made to give a final concentration of 100 ng/μl for use in polymerase chain reaction (PCR).

PCR amplifications were conducted in a total volume of 25 μl using the primers of Crandall and Fitzpatrick (1996), which amplify a 495-bp segment of the 16S gene. Primers and their sequences are 1472 5′-AGATAGAAACCAACCTGG-3′ and 16s–17sub 5′-ATASRGTCTRACCTGCCC-3′. Each reaction contained the following components: 1× PCR buffer, 2.5 mM magnesium chloride, 1.25 mM each dNTP, 1 μM each primer, 0.6 units of *Taq* DNA polymerase (Promega, Madison, WI), and 250 ng of sample DNA. PCR cycling conditions included an initial denaturation step of 2 min at 96°C followed by 45 cycles at 95°C for 30 sec, 41°C for 45 sec, and 72°C for 1 min and 45 sec. A final extension at 72°C for 7 min was then conducted, followed by a soak at 4°C. The PCR products were then run on a 1% agarose gel and the bands excised for sequencing. Before sequencing, the amplified DNA was purified from the gel slices using the GeneClean III kit (Qbiogene,Carlsbad, CA) and concentrated into 10 μl of TLE buffer (see above).

Sequencing reactions were conducted in a total volume of 5 μl using the Big Dye v2 kit from Applied Biosystems (Foster City, CA). Each reaction contained 2 μl of Big Dye ready reaction mix, 0.8 μl of the concentrated GeneClean DNA and 2.2 μl of primer (10 μM). The cycle sequencing protocol followed the manufacturer's recommendations. After amplification, the sequencing products were cleaned using sephadex G-50 fine columns and dried down before running on an ABI 377XL (Applied Biosystems) automated sequencer. The sequences obtained from the automated sequencer were initially corrected and aligned using the program Sequencher ver. 3.1.1 (Gene Codes Corp., Inc., Ann Arbor, MI) and then adjusted, as appropriate, by eye.

Nested Clade Analysis

Geographic (great circle) distances among populations were determined from latitudinal and longitudinal coordinates. Coordinates were determined for each population using an online geographic mapping service (http://www.juggling.org/bin/do/map-find) that determines the position of any point in the contiguous Unites States. These coordinates were then entered into an Excel spread sheet that was used to calculate the great-circle distance between populations (in kilometers). River distances (also in kilometers) were determined by tracing river courses using a digital chartmeter and a 1:200,000 scale topographic map. These river distances are rough estimates because the maps do not take into account elevational changes and because precise tracing of the river courses is not always possible.

It should also be noted that populations sampled from the southern Ozarks (Current and St. Francis Rivers) had the largest pairwise river distances (up to 1983 km) in comparisons made between them and populations sampled from other drainages. This is of note because the river courses followed to obtain these distances went through a large area (i.e., much of northern Arkansas) where *O. luteus* does not occur, so it may not seem realistic to think that population-to-population migrations through these areas are possible, but rather occur (or have occurred in the past) by some other means. However, it also seems that the inclusion of these larger (and possibly biased) river distances may actually be useful in highlighting certain anomalous clades that may be overlooked in a standard geographic distance analysis because of a lack of significance due to small clade and nested clade distances (i.e., leading to an inference of panmixia).

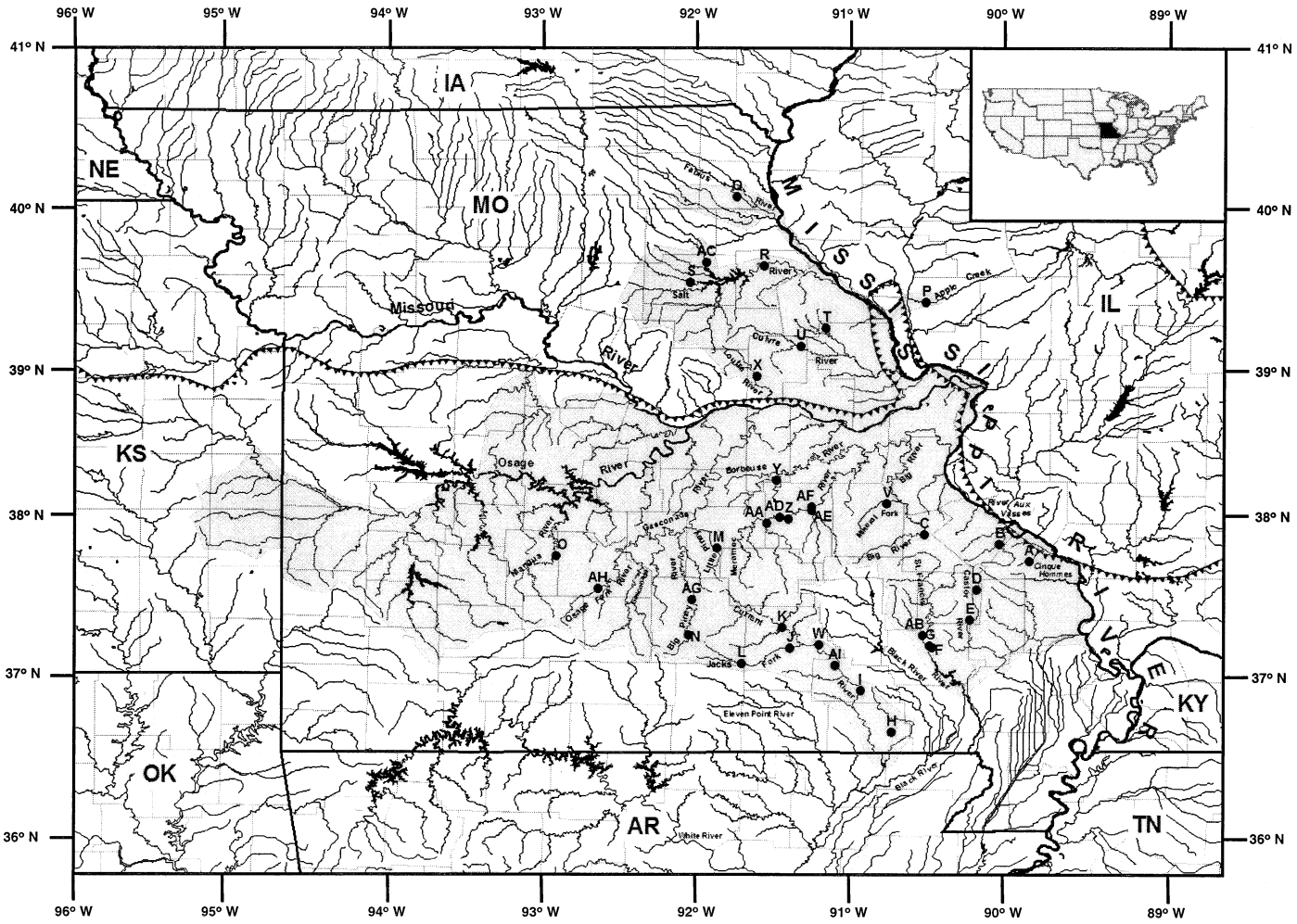


FIG. 3. Map of Missouri (MO) showing the distribution of *Orconectes luteus* (gray area) and the collecting sites (dots) examined in this study. Letters above dots indicate the sampled population names listed in Appendix 2. The line of downward facing triangles indicates the maximum extent of (Illinoian) Pleistocene glacial advances. AR, Arkansas; KY, Kentucky; TN, Tennessee; IL, Illinois; IA, Iowa; NE, Nebraska; KS, Kansas.

The haplotype network was constructed using the program TCS ver. 1.13 (Clement et al. 2000; available from http://inbio.byu.edu/Faculty/kac/crandall_lab/programs.htm) and nesting categories were assigned (see below) following Templeton (1998), Templeton and Sing (1993), and Crandall (1996). Root probabilities were also calculated in the TCS program following the method of Castelloe and Templeton (1994). These probabilities are based on the frequency of a haplotype and the number of mutational connections it has to other (generally lower frequency) haplotypes (Crandall and Templeton 1993).

Ambiguous connections (loops or reticulations) in the haplotype network were resolved using approaches from coalescent theory (see Crandall et al. 1994). In the case of DNA sequence data, this resolution generally involves a comparison of the probabilities of whether a haplotype arose via mutation from either a high- or low- frequency haplotype. In most situations where reticulations occur, this is the comparison being made (high vs. low) and is the case for the current dataset. Coalescent theory would then suggest that, based on these probabilities, the new haplotype arose from the higher-frequency haplotype.

Statistical analyses of geographic associations were conducted using the GeoDis ver. 2.0 program (Posada et al. 2000), which is also available from http://inbio.byu.edu/Faculty/kac/crandall_lab/programs.htm. All statistical analyses in GeoDis were performed using 10,000 (Monte Carlo) replications. Results obtained from GeoDis were then interpreted using the revised inference key of Templeton (1998), available at the above website.

In brief, the nesting process begins by first constructing a haplotype network from a set of aligned gene sequences, or restriction site data, by using the minimum number of mutational connections between haplotypes. Then starting at the tips of the network, the haplotypes are grouped into clades by joining haplotypes (or higher-level clades later in the nesting process) that are one mutational step away from each other. This nesting continues until all haplotypes (or higher-level clades) are grouped into a single nested clade. The NCA method is then used to infer the underlying population pro-

cesses for each clade that demonstrates a significant geographical association (tested using the program GeoDis). This is accomplished by examining two distance measures generated from the population data, the clade distance (D_C) and the nested clade distance (D_N). The clade distance is defined as the average distance of an individual from the geographic center of all individuals within the same nesting clade. In other words, it measures the geographical spread of a clade. The nested clade distance measures how a clade is geographically distributed relative to other clades in the same higher-level nesting category (Templeton et al. 1995; Templeton 2001). In other words, this distance measures the geographic spread of a clade relative to its older, but presumably closest, evolutionary cousins. A distinction is also made between interior and tip (I-T) clades. An interior clade is one that has two or more mutational connections, whereas tip clades only have a single connection. Testing for significantly small or large D_C or D_N distances in each nested clade is then accomplished through Monte Carlo permutations (in GeoDis). Finding a significant departure from simulated randomness leads to the rejection of the null hypothesis of no association between haplotype distributions and geography (i.e., panmixia).

For example, past fragmentation events (a common inference in many studies) will tend to limit the geographical range of clades, which results in significantly small clade distances for both interior and tip clades with a significant restriction in clade distances at higher clade levels (Templeton et al. 1995). This pattern results from a tendency of fragmented populations to accumulate fixed mutational differences after a fragmentation event. Likewise, specific patterns for D_C and D_N are expected for other population processes (see Templeton et al. 1995).

The D_C and D_N distances are generally calculated from information given about the geographic location of each population. For linear (river) distances, the following equations are used to calculate these distances:

$$D_C = \frac{\sum_{i=1}^{K} \frac{n_i(n_i - 1)}{2} \cdot 0 + \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} n_i n_j D_{ij}}{\sum_{i=1}^{K} \frac{n_i(n_i - 1)}{2} + \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} n_i n_j} \quad \text{and} \tag{1}$$

$$D_N = \frac{\sum_{i=1}^{K} \left[\frac{n_i(n_i - 1)}{2} + n_i(N_i - n_i)\right] \cdot 0 + \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} n_i N_j D_{ij}}{\sum_{i=1}^{K} \left[\frac{n_i(n_i - 1)}{2} + n_i(N_i - n_i)\right] + \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} n_i N_j}, \tag{2}$$

where K is the total number of populations sampled, D_{ij} is the user input distance between populations i and j, n_i is the number of copies of the clade x in population i and N_i is the number of copies in population i of the $x + 1$ step clade within which the focal clade is nested (A. Templeton, pers. comm.). These values are automatically calculated in the GeoDis program from a supplied matrix of user-defined distances.

One major drawback of calculating and using river distances, however, is that with a large number of populations in the analysis the number of pairwise distances that must be manually mapped increases dramatically, on the order of

$$\frac{K(K - 1)}{2}, \tag{3}$$

where K is the number of populations. Even for this moderate study of 35 populations, a matrix of 595 pairwise population distances needed to be generated (Appendix 1).

F-Statistics

Traditional *F*-statistics were calculated using the analysis of molecular variance method (AMOVA, Excoffier et al. 1992), which partitions the observed genetic variation into components of within individual, among individuals, and among population differences. Estimates of population pairwise genetic distances were calculated as linearized F_{ST}s (Slatkin 1995) using the Arlequin ver. 2.001 package (Schneider et al. 2000).

Phylogenetic Methods

Phylogenetic analyses were conducted using PAUP* 4.0b10 (Swofford 2000) and only included unique haplotypes. We conducted an unweighted parsimony search using 1000 replicate random stepwise-addition heuristic searches because our dataset contained several informative indels that would be ignored in distance- or likelihood-based analyses. These gap characters were treated as a fifth character state in our analyses. The nodes of the resulting haplotype (gene) tree were then tested for statistical support using the bootstrap (Felsenstein 1985) with 1000 pseudoreplicates. Sequences from two additional species (*Orconectes ozarkae* and *O. punctimanus*), both also collected from and distributed in the Ozarks region, were included as outgroups.

RESULTS

Sequence Data

A total of 393 individuals from 35 different populations were sequenced for a 495-bp region of the mitochondrial 16S gene. Forty-nine sites were found to be variable and resulted in the detection of 39 distinct haplotypes (Genbank accession nos. AF376483–AF376521, Appendix 2). The number of mutational steps among the 39 unique haplotypes ranged from a low of one to a high of 21 (or 0.002–4.2% uncorrected sequence divergence, respectively). This is a surprisingly high level of divergence for a within-species comparison, especially for crayfish (Fetzner and Crandall 2001). The inference of numerous missing haplotypes (i.e., undetected but inferred from single-step mutations) separating regional samples (Fig. 4) suggests a high degree of population differentiation, which contrasts with allozyme studies that normally show low levels of variability among crayfish populations and species (Fetzner 1996; Fetzner et al. 1997). Unique 16S haplotypes were detected in almost every population sampled, and for the most part haplotypes were not shared among regional population groupings (see Fig. 4, Appendix 2). Overall, haplotype diversity within populations was low with 20 of the 35 sampled populations containing only a single 16S haplotype, although the sample sizes for some of these populations were extremely low. On average, however, the Current and Jacks' Fork Rivers contained the highest haplotypic

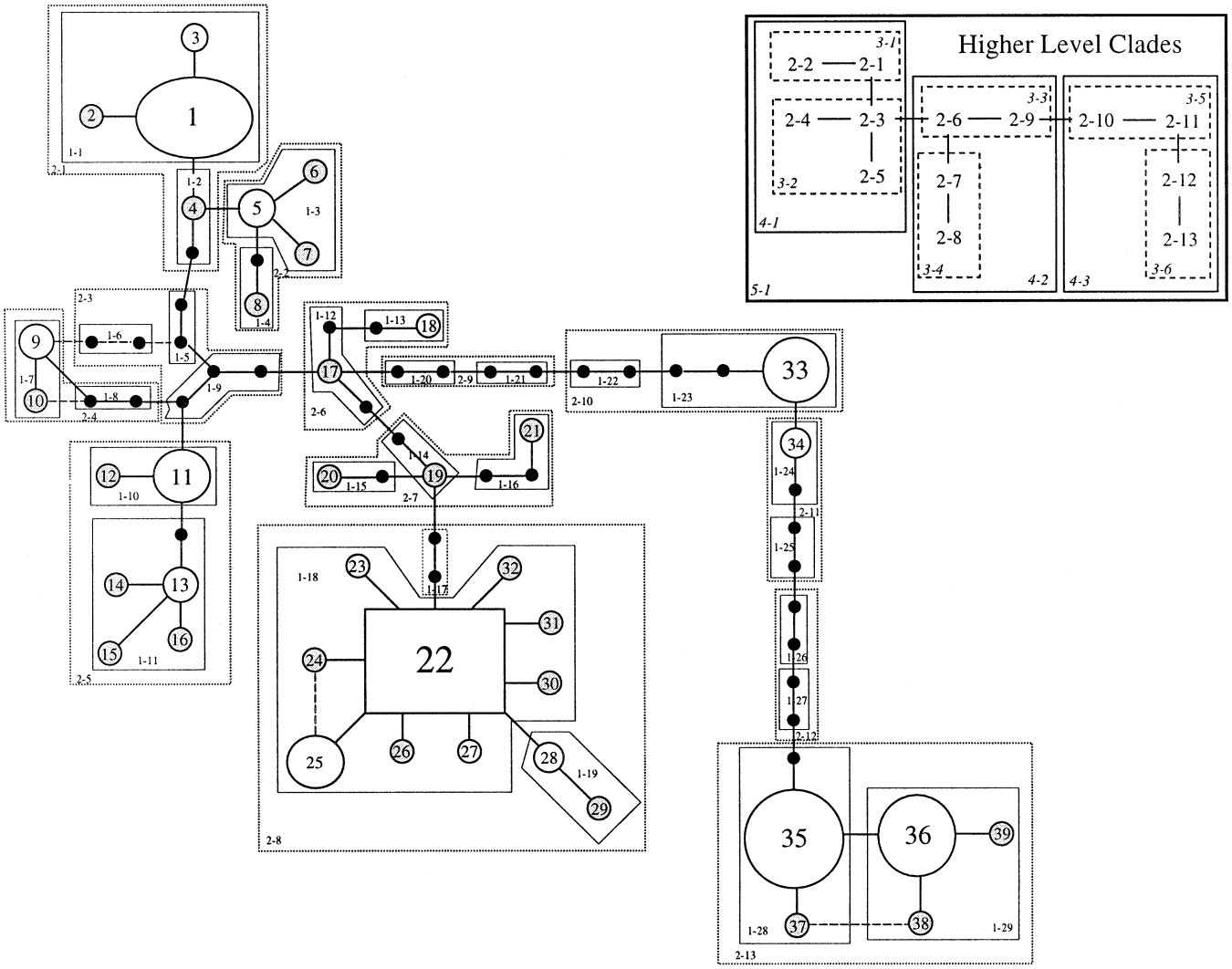


FIG. 4. Haplotype network estimated from the 16S data. Haplotype frequencies are proportional to the area of the associated box or circle. Small black circles are inferred missing haplotypes that were not observed in the data. The large box (haplotype 22) is the haplotype with the largest root probability (Castelloe and Templeton 1994). The large oval (haplotype 1) shows the haplotype found in the highest frequency ($N = 79$). The figure also shows the nesting used to infer the underlying population processes.

diversity, with one population containing up to nine different haplotypes (see Appendix 2). AMOVA estimates of F_{ST} were correspondingly high (overall $F_{ST} = 0.97$) and inferred migration rates among populations were extremely low ($Nm = 0.003$).

Golden Crayfish Phylogeography

When considering mitochondrial DNA haplotypes, the golden crayfish populations appear to be highly subdivided, with little, if any, exchange of migrants among them. Such data suggests that almost every major river drainage in Missouri contains a genetically unique population of this crayfish species. These data also suggest that many of these river populations are old, given that most have accumulated mutational differences and that these mutations have become fixed (or have obtained very high frequencies) in these populations.

However, given the matrilineal inheritance of the mitochondrial genome, it seems possible that any biases in migration rates between the sexes (i.e., higher dispersal distances and rates in males compared to females) may go undetected with this dataset. Some comparisons of this dataset to one containing data from nuclear genes (such as microsatellites) would be helpful in more accurately assessing levels of gene flow among populations and allow for the detection of any bias in migration rates between the two sexes. Such a study is currently being undertaken in the authors' laboratory.

Using the mitochondrial data, several major geographic population groupings were detected and include (1) northeastern Missouri populations along with more southern Mississippi River drainages and the single Illinois population (Fig. 3; populations A-B, P-U, X, AC); (2) the Niangua, Meramec, and Big Rivers (C, O, V, Y-AA, AD-AF); (3) the Osage, Big Piney, Current, and Jack's Fork Rivers (I-N, W, AG, AI); (4) Logan Creek (H); and (5) the St. Francis and

TABLE 1. Comparison of clades with significant geographic associations for both the geographic and linear river distance analyses and the inferences made in each case (see also Figs. 5, 6). Acronyms used in the text follow each inference in parentheses.

Clade	Geographic distances	River distances
1-1	restricted gene flow with isolation by distance (rgf/ibd)	past fragmentation (pf)
1-18	panmixia	inconclusive outcome
2-1	panmixia	restricted gene flow with isolation by distance (rgf/ibd)
2-5	restricted gene flow with some long-distance dispersal (rgf/ldd)	past fragmentation (pf)
2-6	past fragmentation (pf)	panmixia
2-8	contiguous range expansion (cre)	panmixia
2-13	restricted gene flow with isolation by distance (rgf/ibd)	restricted gene flow/dispersal with some long-distance dispersal (rgf/dispersal ldd)
3-1	restricted gene flow with isolation by distance (rgf/ibd)	long-distance colonization (ldc)
3-2	past fragmentation (pf)	past fragmentation (pf)
3-4	past fragmentation (pf)	more sampling needed
4-1	past fragmentation (pf)	past fragmentation (pf)
4-2	past fragmentation (pf)	past fragmentation (pf)
4-3	allopatric fragmentation (af)	allopatric fragmentation (af)
5-1	long-distance colonization (ldc)	long-distance colonization (ldc)

Castor Rivers (D-G). Most of these groups are separated by multiple missing intermediate haplotypes in the network (Fig. 4), suggesting a long evolutionary separation of these populations. The highest root probability was assigned to haplotype 22 ($P = 0.108$), which suggests the Current and Big Piney River drainages, which contain this haplotype, played an important role in the evolutionary history of this species.

Geographic versus River Distance Analyses

For the linear river calculations, population-to-population distances ranged from 2.7 to 1983 (mean = 906.5 km) while standard great circle distances ranged from 2.7 to 388 (mean = 145.8 km; Appendix 1). Comparisons of these distances (Fig. 2) support the use of river distances as the most appropriate for this species, given its affinities for aquatic environments and the significantly large increase in the mean population-to-population distances between the river and geographic analyses (*t*-test: $t = -33.3$, df = 594, $P < 0.0001$).

Overall, the nesting procedure produced nineteen clades that could be tested for a geographical association, and of these, fourteen actually resulted in a significant association (i.e., one or more significant D_C, D_N or I-T values). Both of the distance measures found 12 significant clades, although these 12 clades were not necessarily the same for both methods (Table 1; Figs. 5, 6). Both methods inferred exactly the same events at the higher (3- and four-step) clade levels, with fragmentation events appearing to be a major contributor to the observed pattern of variation at this level (Table 1). The oldest event, however, was detected at the four-step clade level with an inference of long distance colonization, referring to a colonization event from the Current River to Logan Creek and the St. Francis and Castor Rivers.

At the lower zero-step to two-step clade levels, however, the inferences made and even the clades that demonstrated significant associations, depended heavily on the population-to-population distance used in the analysis. Only one of the 10 significant clades at these levels resulted in a similar inference using the two types of distances (clade 3–2, Table 1). Additionally, both methods inferred restricted gene flow for clade 2–13, however, geographic distances also inferred isolation by distance, whereas river distances inferred dispersal with some long-distance dispersal. Clearly, based on these results, the use of linear river distances makes a dramatic difference in the inferences made about the more recent events shaping patterns of genetic variation in these populations. The specific differences for each of these clades are discussed more fully below.

Comparison of Significant Clades

In comparing the two different NCAs, the inferences made for lower-level clades agreed totally for only one of the 10 significant clades (Table 1). Interestingly, the use of river distances resulted in an overall increase in the number of significant values detected per clade ($N = 58$) compared to those from standard geographic distances ($N = 42$) (Table 2). Most of the increase for the linear river analysis was due to a greater proportion of significant nested clade (D_N) and interior-tip (I-T) distances, with many of these being significantly large (Table 2). As a result, the choices made while working through the dichotomous inference key (Templeton 1998, p. 396) were affected. Specifically, the choice made at couplet 3, which states,

> ''Is at least one of the following conditions satisfied?
> a. Are any D_N and/or I-T D_N values significantly reversed from the D_C values?
> b. Do one or more tip clades show significantly large D_N's with the corresponding D_C values being non-significant?
> c. Do one or more interior clades show significantly small D_N's with the corresponding D_C values being non-significant?
> d. Does I-T have a significantly small D_N with the corresponding D_C values being non-significant?
> NO—Go to step 4.
> YES—Go to step 5.''

was influenced by the greater number of significantly large D_N values detected in the river distance analysis (compared to geographic distances). This led to the selection of couplet

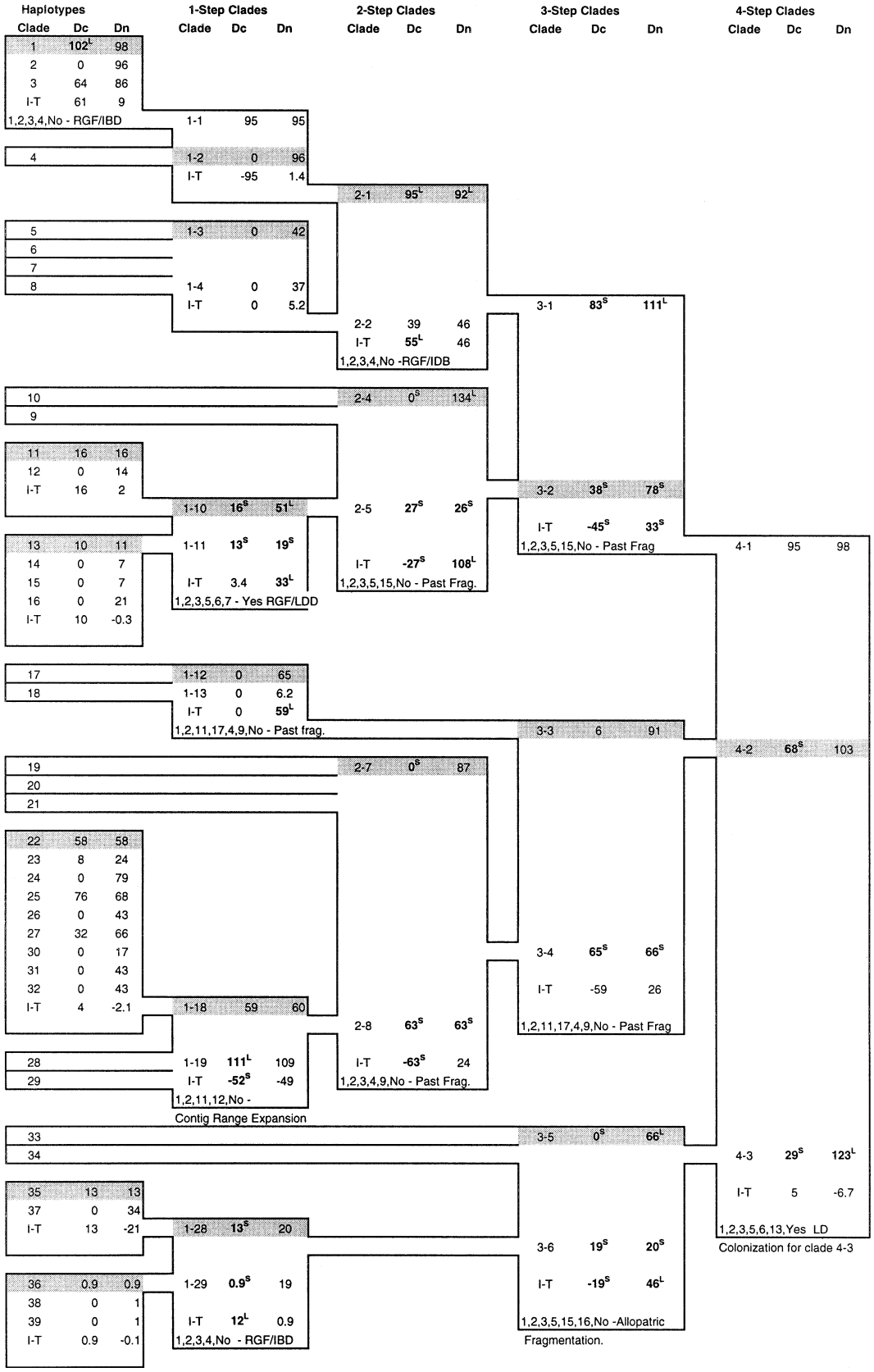


FIG. 5. Results from the nested analysis using geographic distances among populations. Significantly large or small D_C, D_N, and I-T_C or I-T_N values are indicated along with the inference made for clades showing significant associations. Interior clades are indicated by gray stippling.

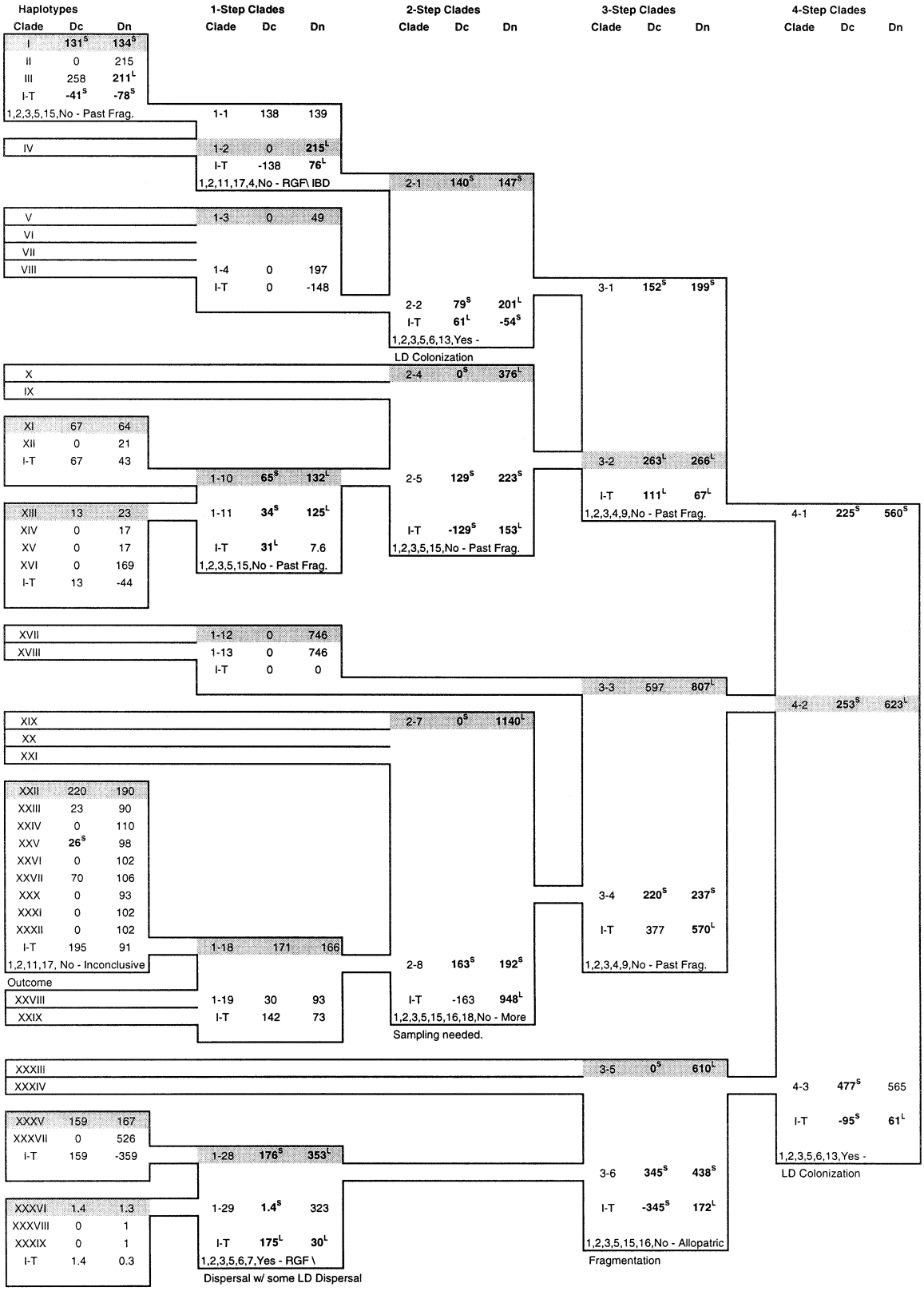


FIG. 6. Results from the nested analysis using river distances among populations. Significantly large or small D_C, D_N, and I-T_C or I-T_N values are indicated along with the inference made for clades showing significant associations. Interior clades are indicated by gray stippling.

TABLE 2. Comparison of D_C, D_N, and I-T values for clades showing a significant association. (S/L), the number of values that were significantly small and significantly large, respectively.

	Geographic (S/L)	River (S/L)
D_C	18 (15/3)	20 (19/1)
D_N	12 (6/6)	20 (8/12)
I-T	12 (6/6)	18 (6/12)
Totals	42 (27/15)	58 (33/25)

5 for many of the clades when using river distances, rather than couplet 4, as was the case in the geographic distance analysis. For this reason, the two analyses resulted in different inferences for the lower-level clades because these two couplets (4 and 5) ultimately led to different areas of the inference key and thus, in almost all cases, inferred different population processes.

If the geographic and river distances are examined by clade (Fig. 7), some general patterns can be seen, however, these patterns tend to be inconsistent in terms of predicting both significant associations and the resulting inference of a population process (i.e., a large gap in population-to-population distances within a clade did not always result in a significant association). For example, in the river analyses, some of the major inferred fragmentation events also show a large break in the population-to-population distances (e.g., Fig. 7C, clade 3-4, vertical axis) as one might expect. However, similar large distances among populations at lower clade levels (Fig. 7A, clade 1-18, vertical axis) were either nonsignificant (suggesting panmixia) or resulted in an inconclusive outcome. The higher nesting levels seem to be less affected because of the larger number of population comparisons made in these groups. Another potential problem at lower clade levels, at least with this dataset, is the presence of many unique population-specific haplotypes. Such a pattern results in a clade distance (D_C) of zero for that haplotype (or clade). This means that the geographic spread of these haplotypes are extremely restricted, so inferences at the lowest (i.e., youngest) clade levels cannot be made, resulting in a failure to reject the null hypothesis of no association (see Figs. 5 and 6, haplotype clades).

Crayfish Population Structure in a Linear Habitat

Unlike many recent NCA studies of freshwater organisms where only one or two major population process were inferred (Turner et al. 2000; Bernatchez 2001; Hurwood and Hughes 2001; Sivasundar et al. 2001b; Fairley et al. 2002; Schultheis et al. 2002), there appear to have been many different processes acting upon golden crayfish populations to produce the currently observed pattern of genetic variation (Table 1). These processes include restricted gene flow with isolation by distance (rgf/ibd; five clades), restricted gene flow with long distance dispersal (rgf/ldd; two clades), contiguous range expansion (cre; one clade), long-distance colonization (ldc; two clades), past fragmentation (pf; six clades), and allopatric fragmentation (af; one clade). See Templeton et al. (1995) for an in-depth discussion of the expected distance patterns associated with these various processes. In addition, one clade for the river distance analysis resulted in an inconclusive outcome due to a lack of sampling in an intermediate area between two populations.

For clade 1-1, the inference from geographic distances was rgf/ibd, whereas river distances suggested a past fragmentation event, either of which seem plausible given the distribution of haplotypes in this clade. It appears that significance in this clade arises when comparing the distribution of haplotype 1 to that of haplotype 2. Haplotype 2 only occurs in the Loutre (X) and Middle Fabius Rivers (Q), whereas haplotype 1 is more widespread geographically. The Loutre (X) and Middle Fabius (Q) Rivers are fairly well separated geographically, so their sharing a single haplotype seems surprising. However, given the limited sampling at these sites, it is possible that other haplotypes may be detected with more thorough sampling, but may also indicate some prior connection between these two drainage systems.

For clade 1-18, the river distance analyses resulted in an inconclusive outcome, due to only a single significantly small D_C value for haplotype 25, whereas geographic-distances analyses failed to detect any significant associations in this clade whatsoever (i.e., panmixia).

The river distance analysis for clade 2-1 inferred rgf/ibd, whereas geographic distances again failed to detect any association. The river distance results stem from the comparison between haplotype 4 and clade 1-1 (haplotypes 1, 2, and 3). Haplotype 4 is only found in Deer Creek (AC) and thus results in a larger than average distance to the higher-level nesting clade. Deer Creek (AC) appears to contain several unique haplotypes when compared to other populations within the Salt River (R, S) drainage system.

For clade 2-5, river distance analyses result in the inference of past fragmentation, whereas geographic distance analyses suggest rgf/ldd. Comparisons for this clade include those among the Big (C, V), Bourbeuse (Y), and Meramec River (Z, AA, AD, AE, AF) systems and the Niangua River (O). The distances between these groups of populations are quite large, and these populations do not share any haplotypes in common. In this case, given the large distances involved, fragmentation seems the most likely event to describe this pattern. Several populations intermediate to these have yet to be sampled, and the addition of these populations to the dataset will shed additional light on the historical process acting on this group of populations.

For clade 2-6, geographic distance analyses suggest past fragmentation, whereas river distance analyses did not detect any significant association. This clade includes two peculiar haplotypes, one from the Harris Branch of the Big River (C) and the other from the St. Francis River (AB). These haplotypes are strange because they are quite distantly related from other haplotypes seen from the same population or from nearby populations in the same drainage. This is especially true for the St. Francis haplotypes. Additional sampling from this area may shed more light on the distribution of these haplotypes.

For clade 2-8, geographic distance analyses suggest contiguous range expansion, whereas river distance analyses again found no significant association. The pattern detected here is similar to that of clade 1-18 for river distances, with the expansion of clade 1-18 haplotypes from the Big Piney

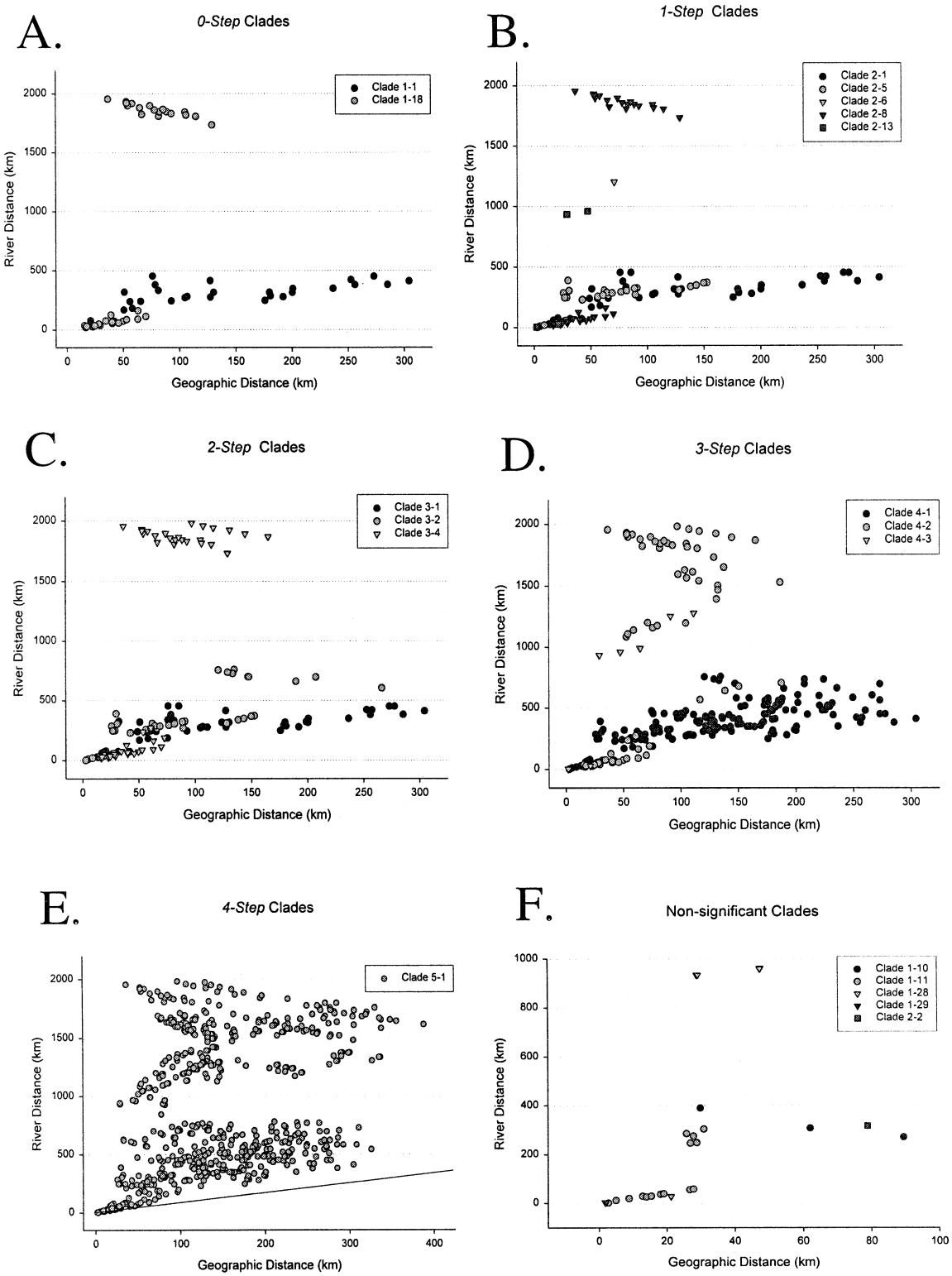


FIG. 7. Plots of geographic and river distances by clade and clade level. (A) 0-step clades; (B) 1-step clades; (C) 2-step clades; (D) 3-step clades; (E) 4-step clades; and (F) the five clades that showed no significant associations.

(N, AG) into the Current River (H, I, J, K, L, W, AI; or vice versa).

For clade 2-13, both analyses inferred rgf but geographic distances suggested ibd, whereas river distances suggested rgf/dispersal with some long-distance dispersal. This clade involves the St. Francis (F, G) and Castor River (D, E) populations. These populations are quite close geographically, however, the river distances between them are quite large.

Given the distribution of this species, it seems highly likely that these populations have exchanged migrants by means other than by following river courses (see Discussion).

For clade 3-1, the two analyses again came up with entirely different inferences. Geographic distance analyses suggested rgf/idb, whereas river distances inferred long-distance colonization. The clades involved in this comparison include clades 2-1 and 2-2. Clade 2-2 contains four haplotypes that are found in two populations, Deer Creek (AC), and West Fork Cuivre River (U), whereas the haplotypes in clade 2-1 are much more geographically widespread.

For clade 3-4, the geographic distance analyses suggest past fragmentation as reason for the observed patterns between clades 2-7 and 2-8. Clade 2-7 includes three haplotypes sampled from the Osage Fork River (AH) and clade 2-8 includes the Current River (H, I, J, K, L, W, AI) and Big Piney (N, AG) populations. The river distance analyses, however, resulted in an inconclusive outcome due to a lack of data from populations intermediate between these two clades. Additional sampling from these missing populations will be needed before we can accurately assess the inferences made for this clade.

Discussion

Geographic versus River Distance Analyses

The two different NCAs examined here arrived at quite different conclusions for lower-level clades (see Table 1; Figs. 5, 6) but suggested the same processes for the higher-level clades. These results suggest that this choice of distances used in the NCA can strongly influence the outcome of inferences made. The river distances exhibit higher levels of variation at these low clade levels when compared to the geographic distances, and this increased level of separation among population groups (Fig. 7) may account for the differences detected between the two methods. Additional studies will need to be made to see how widespread this phenomenon is. For example, in the recent study by Turner et al. (2000), the authors conducted an NCA on a freshwater mussel species from Arkansas. They used river distances in their analysis, but they did not compare their results to the standard geographic distance analysis.

The results presented here suggest that it is much easier to consistently infer older events than it is to infer the more recent ones. This is related to the larger distances involved between the older clade levels for both methods. Presumably, at these higher clade levels, the two distances have more similar magnitudes and thus are able to infer similar processes, whereas a denser sampling scheme is needed to infer the younger events (i.e., to discriminate among different alternative processes).

Given these results, one is left to decide which distance measure is most appropriate for the organism at hand. In fact, a combination of the two may be useful in explaining historical and contemporary processes. For example, populations from the Current (H, I, K, W, AI) and Jack's Fork (J, L) Rivers have most likely exchanged migrants with the Big Piney River (N, AG) system as indicated by their sharing haplotype 22 at a high frequency. In this case it seems that this exchange of individuals occurred by means other than by following (downstream) river courses that ultimately traverse areas where the species does not occur. However, the headwaters of these systems are in very close proximity geographically, and a transfer of individuals from these areas seems highly probable. This transfer may have occurred either through a change in stream drainage patterns caused by glacial advances during the mid to late Pleistocene or possibly as an overland migration event. In contrast, most of the other populations (drainages) probably exchange migrants by following actual river courses. The exchange of migrants among drainages (and subdrainages in some cases) appears to be rare, as evidenced by the complete fixation of alternate haplotypes in many of these stream systems (see Appendix 2).

So, which distance measure should be used in a NCA? For entirely aquatic organisms (fishes, mussels, etc.) the use of river distances would be most appropriate because these are the actual distances these organisms must traverse to exchange genetic material. Crayfishes are somewhat unique in this regard because of their potential for migrations over land (Lodge et al. 2000). To our knowledge, however, there have not been any documented cases of this phenomenon in the golden crayfish.

Golden Crayfish Population Structure

There seems to be sufficient levels of variation in the mitochondrial 16S gene for examining population structure in the golden crayfish and other crayfish species. It also appears that past fragmentation has played an important role in the early history of this species and helped to define the current population structure of the golden crayfish in the Ozarks of Missouri. Such fragmentation events are likely a result of Pleistocene glaciation events, which altered pre-Pleistocene river drainage patterns in the Ozarks region (Mayden 1988; Crandall and Templeton 1999). Given the high levels of divergence among the regional samples of this species, it may suggest the presence of multiple refugia for this species during the Pleistocene. It is actually quite surprising how closely this species' distribution matches the southern extent of the glacial maximum (see Fig. 3). Given that the inferred root haplotype (22) from the network resides in the Current River populations, it is possible that this may have been one such refugium for this species. Long-distance colonization events (inferred by clade 5-1) would then have occurred into the St. Francis region as well as the Big Piney and possibly the Meramec River. Another potential refugium may have been in northeastern Missouri in and around the lower Salt River or lower Cuivre River drainages. Subsequent invasion into the Mississippi River after the retreat of the ice sheet would give these populations access to streams both further up- and downriver and allow new populations to establish themselves in the Fabius River and Cinque Hommes and River Aux Vases Creeks, and even allow for the invasion of Illinois.

Overall, there is a general pattern of within-drainage haplotype uniformity in this species, suggesting high levels of gene exchange among different populations within drainages. However, among drainages the picture is completely different. It appears that each drainage system contains a unique haplotype or set of haplotypes, at least in all but a few cases.

This suggests that the transfer of individuals among drainages is a rare event, but does occasionally occur.

The data presented in this paper suggest a relatively recent connection between the Current River watershed and the Big Piney and Little Piney River systems. Indeed, these watersheds are the only ones in our sample that shared mtDNA haplotypes at high frequencies. Clearly, some level of interconnection between these systems must have been present in the recent past, although the potential role of overland migration events cannot be ruled out. The other watersheds that share haplotypes, but at lower frequencies, are the St. Francis and Castor Rivers and the Harris Branch of the Big River and Cinque Hommes Creek. For the former, there seems to have been a connection between the lower St. Francis and the Castor River that allowed for the transfer of individuals carrying haplotype 35. In the latter case, only a single individual from Cinque Hommes was detected with the most common Harris Branch haplotype.

Potential Means of Migration

Unlike many other aquatic organisms, crayfish have several different means by which they can migrate between or establish new populations. Crayfishes readily disperse along watercourses (Lodge et al. 2000) or within watersheds. Historically, changes in river drainage patterns during glacial episodes in the Ozarks may have played a role in the current distribution of aquatic organisms (Mayden 1988; Crandall and Templeton 1999). The movement of major stream courses from one area to another may have been mediated through stream capture events. Such changes would have stranded residents in the old streambed and simultaneously opened potentially new habitat to species inhabiting the new river drainage area. The populations occurring in the old stretch of the river may then be isolated and allow for the accumulation of mutations among the populations.

One such case in the present dataset may be exemplified by the St. Francis River populations (F, G, and AB; Appendix 2). These populations are only separated by about 12 river miles but differ drastically in their haplotypes. Interestingly, the Bounds and Frazier Creek populations are more closely related to populations in the Castor River system than they are to the other St. Francis River population (Fig. 4, Appendix 2). This haplotypic difference among St. Francis populations may actually record a stream capture event that occurred sometime in the past, most likely when the St. Francis drainage switched from being a direct tributary to the Mississippi River to its current status as a tributary to the White River in Arkansas.

Another possible means of migration has also occasionally been observed. Crayfish can and will leave their aquatic environment and traverse land to new areas, usually under extremely humid conditions (Lodge et al. 2000). The frequency of such events is not known, and some species may do this more than others. A related method of migration may be when crayfish inhabit headwater stream areas (which are often temporary or semipermanent). These areas are often in close proximity to the headwaters of other streams (especially in the Ozarks area). When these areas are in the process of drying out, it may be a stimulus for the inhabitants to search for new areas, and because headwater areas of other streams are in close proximity, the crayfish may be able to establish themselves in new areas, if they are able to survive a journey across land.

A third potential method for establishing new populations is large floods, especially in the areas of the Mississippi flood plain, which may help promote the establishment of new populations by allowing crayfish (and other organisms) to traverse areas that are otherwise dry land under normal conditions. Floods may also promote the mixing of populations by washing individuals downstream under heavy stream currents. For example, flooding events have been implicated as a major factor for the spread for several exotic species in the United States, including the zebra mussel (*Dreissena polymorpha*) and several plant species.

A final means by which crayfish may establish themselves in new areas is through human-aided or ''bait bucket'' introductions. Crayfish are used by fishermen as fishing bait, and they can inadvertently transport members of different populations or species from one river drainage to another (Eng and Daniels 1982; Ludwig and Leitch 1996). These inadvertent human introductions can cause serious problems, not only for vulnerable native species but also for scientists trying to unravel population relationships. In response to these introductions, many states are enacting strict regulations to prevent the transfer and/or introduction of exotic crayfishes and are banning their use as bait by fishermen. We should note that the golden crayfish has not been sold as bait in the state of Missouri, so we do not expect complicated population patterns to arise from human-aided transfers and, in fact, do not see any evidence of this in our dataset.

Biogeography of the Ozarks

The Ozarks region is well known for its cool, clear streams and an abundant freshwater fauna. In general, this area consists of uplifted regions in southern Missouri and northern Arkansas, but also extends short distances into Illinois and Oklahoma. The Ozarks Plateau Region covers an area of about 40,000 square miles and is bounded mainly by the Mississippi, Missouri, and Arkansas Rivers. In southeastern Missouri, the upland areas are drained by tributaries flowing north to the Missouri River (e.g., Osage and Gasconade), to the east by tributaries to the Mississippi River (Meramec and others), and the south by tributaries to the White and Arkansas Rivers (Black, Current, Eleven Point, and St. Francis). The drainage patterns in this region have been greatly influenced by glacial events that occurred during the Pleistocene (Thornbury 1965; Mayden 1985; Robison 1986; Mayden 1988).

In Missouri, the maximum glacial extent reached as far south as the Missouri River (or a bit further in some regions; see Fig. 3). During these glacial events, much of the Ozarks region contained tundra and boreal forests that are now characteristic of more northern climates today. Climactic conditions at that time were such that levels of precipitation were reduced. This reduction in available water led to streams that were both smaller and ultimately less turbid due to reduced levels of glacial till being carried by these streams. These

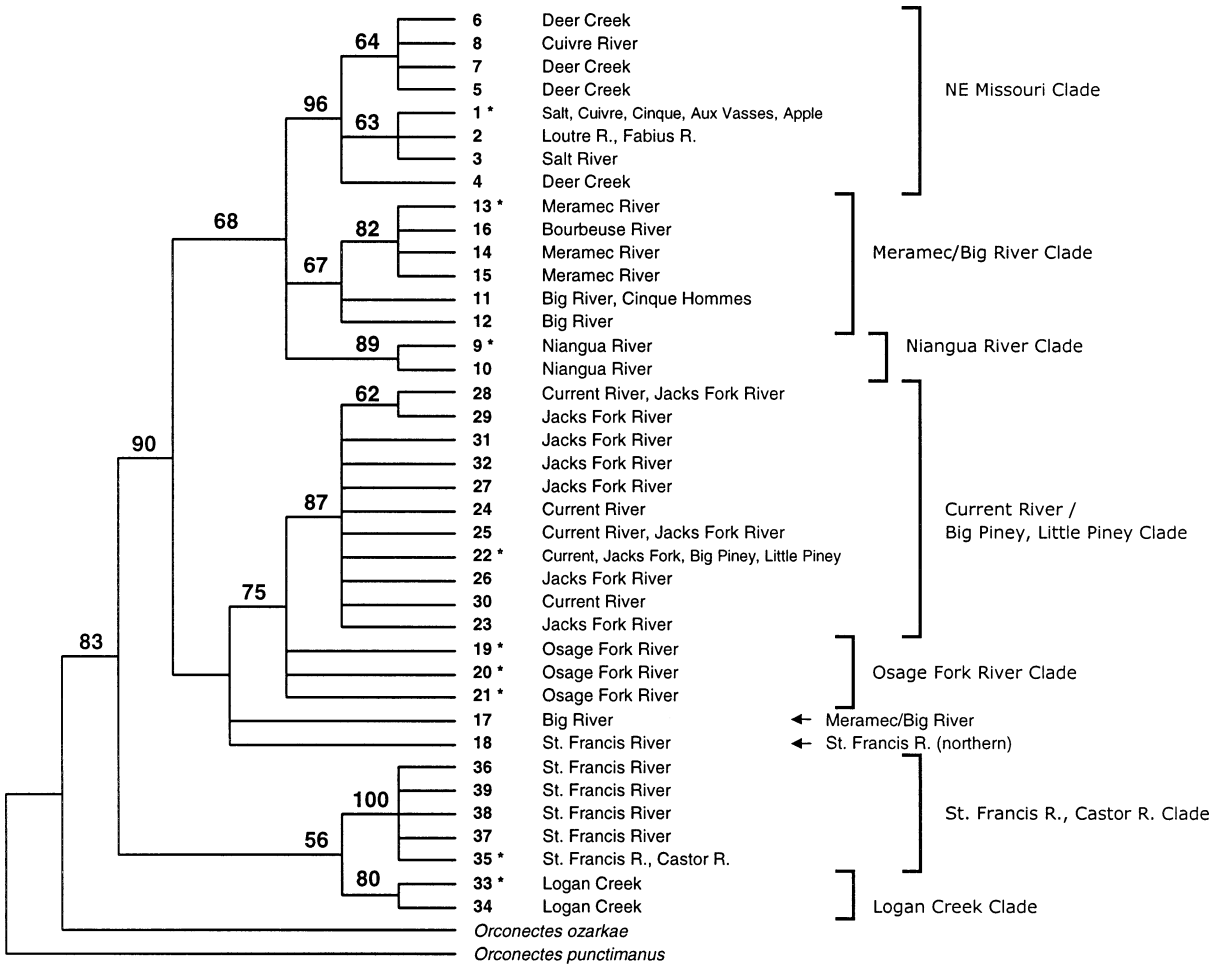


FIG. 8. Strict consensus of 60 most parsimonious trees generated using 1000 random addition sequences and treating gaps as a fifth state. Numbers at branch tips represent the 16S haplotypes detected and are followed by the geographic locations where they were found. An asterisk indicates the haplotype in each major clade with either the highest frequency or the broadest geographic spread. Numbers above the branches represent bootstrap support (1000 replicates). Tree length = 101, CI = 0.782, RI = 0.914.

conditions then facilitated the expansion of highland organisms, such as crayfish, adapted to clear, cool streams.

The lower Missouri River, which consists of the region extending from Kansas City downstream to it confluence with the Mississippi River at St. Louis, is thought to have undergone relatively little change in its drainage pattern since preglacial times (Thornbury 1965; Robison 1986). However, this is not the case for upper reaches of this enormous river system, which appears to have been captured from a preglacial Hudson Bay drainage system. The Meramec River system also appears to have been affected little by the advance and retreat of glaciers.

While tributaries of the northern Ozarks are thought to have changed little as a result of Pleistocene glacial events, the southern tributaries (Current, St. Francis, Eleven Point, Black, and White) are an entirely different story. In pre-Pleistocene times, the Mississippi River is thought to have occupied the channel currently containing the Black, White, and Cache Rivers (Robison 1986). In addition, the Ohio River is thought to have occupied the current St. Francis River channel and the Tennessee River occupied the current channel of the Mississippi (Robison 1986). These three major river systems joined together much farther south (near Helena, Arkansas) than they do currently. During the Pleistocene, the Mississippi River was diverted into the Ohio channel near Thebes, Illinois. Later, the Ohio River joined with the Tennessee near its current location.

These changes in drainage patterns had a significant effect on the indigenous aquatic fauna. As a result, previously contiguous habitat was dissected and routes open to dispersal were drastically altered. Zoogeographic studies of aquatic species (mainly fishes) from this region are informative and suggests the possibility of past stream connections (Pflieger 1971; Cross et al. 1986).

Past connections between different Ozark drainages can be hypothesized based on a variety of species distributional data and include connections between the White and Gasconade Rivers and the St. Francis, Black, and White Rivers (but these are complicated due to the previous location of the Mississippi River in this region). We can add to this list a connection between the Current/Jack's Fork Rivers and the Big and Little Piney Rivers (Gasconade drainage).

Several authors have examined phylogeographic patterns of a variety of different freshwater species from the Central

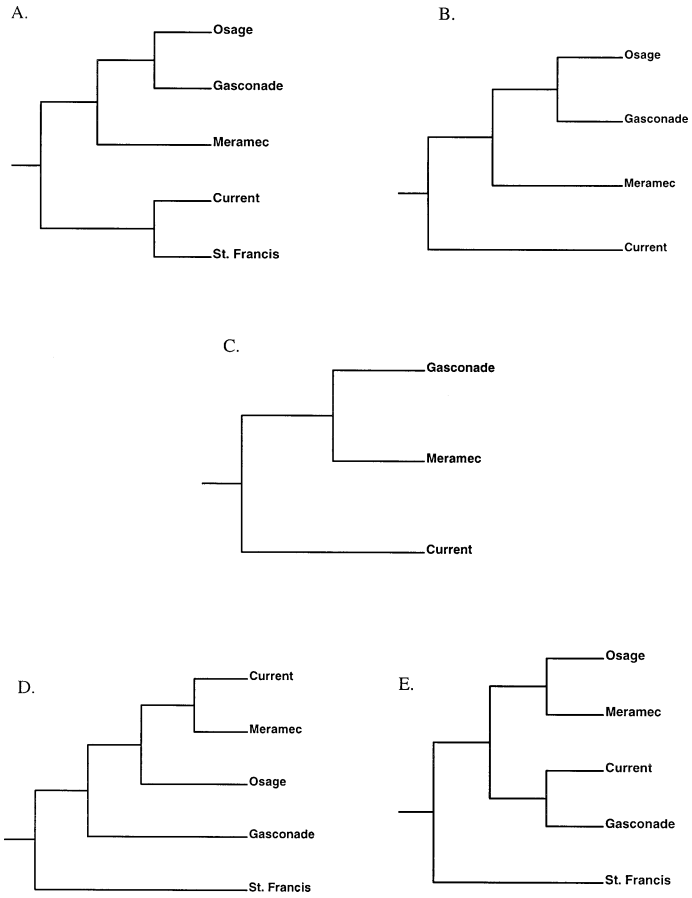


FIG. 9. Area cladograms depicting relationships among the different Ozark river drainages using data from (A) numerous freshwater fish species (Mayden 1988); (B) the hellbender, *Cryptobranchus alleganiensis* (Routman et al. 1994); (C) the slender madtom, *Noturus exilis* (Hardy et al. 2002); (D) crayfish genus *Orconectes* (Crandall and Templeton 1999); and (E) the golden crayfish, *Orconectes luteus* (this study).

Highlands (includes the Ozark, Ouachita, and Eastern Highland regions; Mayden 1988; Routman 1993; Routman et al. 1994; Crandall and Templeton 1999; Turner et al. 2000; Hardy et al. 2002). However, rather than focusing on one particular region, such as the Ozarks, these studies have generally been on a much larger geographic scale (Central Highlands). Therefore, comparing biogeographic patterns across studies is a bit more difficult because the geographic resolution of previous studies was not at the same scale (i.e., they only include major drainages, not subdrainages). However, where possible, we compare our results with theirs below.

In comparing our results of relationships among major Ozark river systems (Fig. 8) to those obtained by other investigators, we were only able to directly compare at most five major drainages, and these include the Gasconade (in our study this includes the Little Piney, Big Piney, and Osage Fork), the Osage (Niangua), the Meramec (Meramec, Big River, Mineral Fork, and Bourbeuse), Current (Current and Jack's Fork), and St. Francis Rivers. In general, the results obtained by Mayden (1988), Routman et al. (1994), and Hardy et al. (2002) were congruent with one another when only these Ozark stream relationships were considered, but differed quite drastically with results obtained by Crandall and Templeton (1999) for other crayfishes (Fig. 9D). Our results from this study also differed from those arrived at previously (see below). The Mayden (1988) fish dataset includes all five drainages, and groups the (Osage, Gasconade) Meramec, which is then sister to a (Current, St. Francis) group (Fig. 9A). The Routman et al. (1994) hellbender dataset (Fig. 9B) recovers the same topology, but does not include data from the St. Francis, while the Hardy et al. (2002) madtom dataset did not sample the Osage (Fig. 9C). Essentially, these three studies support a split between the northern- and southern-flowing Ozark river systems (Mayden 1988).

Drainage relationships inferred from crayfishes are quite different from those obtained in the hellbender and fish examples above. Crandall and Templeton (1999) recovered the following set of relationships among comparable Ozark drainages: ((((Current, Meramec) Osage) Gasconade) St. Francis) (Fig. 9D). On the other hand, our current data suggests the following set of relationships among the same drainages: (((Meramec, Osage) (Current, Gasconade)) St. Francis) (Fig. 9E). The results for these two studies are quite similar and demonstrate the distinctiveness of the St. Francis River drainage. The main uncertainty between these two studies is in the relationships of the Current River. As stated above, this river system appears to have played an important role, possibly as a refugium, with several colonization events occurring from here to other drainages in the region. Therefore, it is not surprising that these drainages show varying affinities in studies involving different species.

It is also quite interesting to note that *O. luteus* is apparently absent in the upper Black River Drainage. If the Black, Current, and St. Francis Rivers joined directly to the Mississippi River in pre-Pleistocene times, and this species occurs in both the Current and St. Francis drainages, it seems odd that the species is absent from the Black. Other species of crayfish occur there and the transfer of crayfish species from the Black to the St. Francis is known to have occurred, even in recent times. So it seems possible that the reverse transfer is equally likely. Clearly, based on the genetic data, this region has had a dynamic history.

The St. Francis drainage also displays an interesting set of relationships. Of the three populations sampled from this river system, the two southernmost sites (F, G) are only about 2 miles apart, while the third, more northern population (AB) is about 10 miles to the north. The interesting point here is that the northern population is quite distinct from the two just a relatively short distance to the south. In fact, the northern population haplotype (18) falls into an altogether different clade (see Fig. 8) and is only distantly related to the southern St. Francis haplotypes. In addition, the presence of another set of very distinct haplotypes in the Logan Creek population (H) tends to highlight this area of southeastern Missouri as being zoogeographically important. Clearly some event has altered the genetic patterns seen in this area, and this may be related to the river drainage patterns shifting from direct tributaries to the Mississippi during the Pleistocene to their present-day pattern. Further study of the fauna in this region of the Ozarks is needed before any real concrete patterns of drainage relationships can be corroborated with those displayed by the crayfish species examined in this study.

Conclusions

It is a very exciting time in population genetics, as new methods become available that merge both the spatial and temporal aspects of genetic variation. The NCA procedure is one of these new methods that holds great promise for studying genetic variation at the population level.

The inclusion of linear river distances into the nested analysis, in place of geographic distances, appears to produce a significant increase in the number of clade, nested clade, and interior-tip distances. The river distances also affected the detection of significant clades and the inferences made for those clades. We suggest that those researchers working with riverine, riparian, or coastal species should conduct both types of nested analyses to compare the inferences made from each. This process may help identify areas where added sampling may be needed or help identify clades that need to be looked at in greater detail. In the case presented here, the comparison of results from both methods tended to highlight cross-drainage transfers of haplotypes.

An Excel spreadsheet file containing the matrices of geographic and river distances used in this study is available upon request from the first author.

Acknowledgments

We thank D. Posada and A. Templeton for the derivations of the equations for calculating D_C and D_N values for river distances and allowing us to include them here. This work was supported by a Doctoral Dissertation Improvement Grant (DEB 0073155) to KAC and JWF and by grant IBN 9702338 to KAC, both from the National Science Foundation. We also thank the Missouri Department of Conservation for supplying JWF with collecting permits.

Literature Cited

Bernatchez, L. 2001. The evolutionary history of brown trout (*Salmo trutta* L.) inferred from phylogeographic, nested clade, and mismatch analyses of mitochondrial DNA variation. Evolution 55: 351–379.

Carbone, I., and L. M. Kohn. 2001. A microbial population-species interface: nested cladistic and coalescent inference with multilocus data. Mol. Ecol. 10:947–964.

Castelloe, J., and A. R. Templeton. 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. Mol. Phylogenet. Evol. 3:102–113.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9:1657–1659.

Crandall, K. A. 1996. Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. Mol. Biol. Evol. 13:115–131.

Crandall, K. A., and J. F. Fitzpatrick, Jr. 1996. Crayfish molecular systematics: using a combination of procedures to estimate phylogeny. Syst. Biol. 45:1–26.

Crandall, K. A., and A. R. Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959–969.

———. 1999. The zoogeography and centers of origin of the crayfish subgenus *Procericambarus* (Decapoda: Cambaridae). Evolution 53:123–134.

Crandall, K. A., A. R. Templeton, and C. F. Sing. 1994. Intraspecific phylogenetics: problems and solutions. Pp. 273–297 *in* R. W. Scotland, D. J. Siebert, and D. M. Williams, eds. Models in phylogeny reconstruction. Clarendon Press, Oxford, U.K.

Crandall, K. A., J. W. Fetzner, Jr., S. H. Lawler, M. Kinnersley, and C. M. Austin. 1999. Phylogenetic relationships among the Australian and New Zealand genera of freshwater crayfishes (Decapoda: Parastacidae). Aust. J. Zool. 47:199–214.

Cross, F. B., R. L. Mayden, and J. D. Stewart. 1986. Fishes in the Western Mississippi drainage. Pp. 363–412 *in* C. H. Hocutt and E. O. Wiley, eds. The zoogeography of North American freshwater fishes. John Wiley and Sons, New York.

Cruzan, M. B., and A. R. Templeton. 2000. Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. Trends Ecol. Evol. 15:491–496.

Dumolin-Lapegue, S., B. Demesure, S. Fineschi, V. Le Corre, and R. Petit. 1997. Phylogeographic structure of white oaks throughout the European continent. Genetics 146:1475–1487.

Emerson, B. C., E. Paradis, and C. Thébaud. 2001. Revealing the demographic histories of species using DNA sequences. Trends Ecol. Evol. 16:707–716.

Eng, L. L., and R. A. Daniels. 1982. Life history, distribution, and status of *Pacifastacus fortis* (Decapoda: Astacidae). California Fish and Game 68:197–212.

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.

Fairley, T. L., M. M. Povoa, and J. E. Conn. 2002. Evaluation of the Amazon River delta as a barrier to gene flow for the regional malaria vector, *Anopheles aquasalis* (Diptera: Culicidae) in northeastern Brazil. J. Med. Entomol. 39:861–869.

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

Fetzner, J. W., Jr. 1996. Biochemical systematics and evolution of the crayfish genus *Orconectes* (Decapoda: Cambaridae). J. Crust. Biol. 16:111–141.

Fetzner, J. W., Jr.,and K. A. Crandall. 2001. Genetic variation. Pp. 291–326 *in* D. M. Holdich, ed. Biology of freshwater crayfish. Blackwell Scientific, Oxford, U.K.

Fetzner, J. W., Jr., R. J. Sheehan, and L. W. Seeb. 1997. Genetic implications of broodstock selection for crayfish aquaculture in the Midwestern United States. Aquaculture 154:39–55.

Hardy, M. E., J. M. Grady, and E. J. Routman. 2002. Intraspecific phylogeography of the slender madtom: the complex evolutionary history of the Central Highlands of the United States. Mol. Ecol. 11:2393–2403.

Hurwood, D. A., and J. M. Hughes. 2001. Nested clade analysis of the freshwater shrimp, *Caridina zebra* (Decapoda: Atyidae), from north-eastern Australia. Mol. Ecol. 10:113–125.

Knowles, L. L., and W. P. Maddison. 2002. Statistical phylogeography. Mol. Ecol. 11:2623–2635.

Lodge, D. M., C. A. Taylor, D. M. Holdich, and J. Skurdal. 2000. Nonindigenous crayfishes threaten North American freshwater biodiversity. Fisheries 25:7–20.

Ludwig, H. R., Jr.,and J. A. Leitch. 1996. Inter-basin transfer of aquatic biota via anglers' bait buckets. Fisheries 21:14–18.

Mayden, R. L. 1985. Biogeography of Ouachita Highland fishes. Southw. Naturalist 30:195–211.

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

Pflieger, W. L. 1971. A distributional study of Missouri fishes. Univ. Kans. Publ. Mus. Nat. Hist. 20:226–570.

Posada, D., K. A. Crandall, and A. R. Templeton. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Mol. Ecol. 9:487–488.

Pybus, O. G., A. Rambaut, and P. H. Harvey. 2000. An integrated framework for the inference of viral population history from reconstructed genealogies. Genetics 155:1429–1437.

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

Rogers, A. R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9:552–569.

Routman, E. J. 1993. Mitochondrial DNA variation in *Cryptobranchus alleganiensis*, a salamander with extremely low allozyme diversity. Copeia 1993:407–416.

Routman, E. J., R. Wu, and A. R. Templeton. 1994. Parsimony, molecular evolution, and biogeography: the case of the North American giant salamander. Evolution 48:1799–1809.

Schneider, S., D. Roessli, and L. Excofier. 2000. Arlequin: a software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, Univ. of Geneva.

Schultheis, A. S., L. A. Weigt, and A. C. Hendricks. 2002. Gene flow, dispersal, and nested clade analysis among populations of the stonefly *Peltoperla tarteri* in the southern Appalachians. Mol. Ecol. 11:317–327.

Sivasundar, A., E. Bermingham, and G. Ortí. 2001a. Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. Mol. Ecol. 10:407–417.

———. 2001b. Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American rivers. Mol. Ecol. 10:407–417.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462.

Swofford, D. L. 2000. PAUP*; phylogenetic analysis using parsimony (* and other methods). Sinauer Associates, Sunderland, MA.

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

———. 2001. Using phylogeographic analyses of gene trees to test species status and processes. Mol. Ecol. 10:779–791.

Templeton, A. R., and C. F. Sing. 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. Genetics 134:659–669.

Templeton, A. R., E. Boerwinkle, and C. F. Sing. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. Genetics 117:343–351.

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.

Thornbury, W. D. 1965. Regional geomorphology of the United States. John Wiley and Sons, New York.

Turner, T. F., J. C. Trexler, J. L. Harris, and J. L. Haynes. 2000. Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. Genetics 154: 777–785.

Wetzel, J. E., and W. J. Poly. 2000. Range expansion of golden crayfish, *Orconectes luteus*, in the Mississippi River. Journal of Freshwater Ecology 15:281–282.

Corresponding Editor: K. Shaw

APPENDIX 1. Table of pairwise population to population distances (km). Geographic (GPS) distances above diagonal and river distances below diagonal.

No.	Population name	*N*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	Cinque Hommes Creek	24	----	21	62	36	53	82	82	144	132	148	147	178	176	199	286	192	304	256	273	200	199	89	132	201	153	138	150	80	278	143	127	128	192	244	132
2	River Aux Vases Creek	30	79	----	43	35	56	82	82	146	130	139	136	168	159	187	249	176	285	237	253	181	179	69	124	180	133	120	131	78	257	124	108	109	177	228	126
3	Harris Branch	21	307	278	----	49	66	80	79	141	116	111	104	138	117	151	207	164	260	210	221	156	150	30	98	146	91	77	89	71	227	82	65	66	139	187	105
4	Cape Creek	35	212	248	476	----	21	48	47	111	97	113	113	143	149	166	238	206	308	258	270	204	199	79	97	195	136	117	127	44	277	122	109	110	161	213	96
5	Castor River	15	185	220	449	27	----	29	29	91	80	104	106	133	151	159	238	226	326	276	285	221	215	94	87	209	146	124	133	29	292	129	118	120	157	211	83
6	Bounds Creek	24	929	966	1194	961	933	----	2	64	51	81	86	109	140	138	222	244	337	286	292	234	226	105	64	216	147	122	128	10	300	126	118	121	140	194	57
7	Frazier Creek	29	927	966	1193	959	933	3	----	65	51	79	85	107	138	137	221	243	335	285	290	232	224	103	63	214	145	120	126	8	299	124	117	119	138	192	55
8	Logan Creek	31	1128	1167	1511	1278	1252	993	991	----	34	83	97	98	163	134	227	305	388	338	338	288	278	161	75	263	189	162	164	71	348	165	163	166	147	195	57
9	Cave Spring Creek	12	1159	1194	1542	1307	1279	1020	1019	85	----	50	63	70	129	105	197	277	356	305	304	257	246	133	40	230	156	128	129	53	315	131	130	133	115	165	23
10	Jacks Fork River - SR19	29	1223	1268	1614	1381	1353	1094	1093	159	74	----	15	29	82	58	148	256	321	271	265	227	214	116	17	194	119	92	89	76	277	93	98	101	65	117	28
11	Current River - Round Spring	29	1239	1286	1630	1397	1371	1112	1110	177	90	----	----	34	67	53	137	244	306	256	250	213	200	106	24	178	104	77	74	80	262	78	84	87	54	108	40
12	Jacks Fork River - SR17	2	1262	1308	1653	1419	1392	1135	1133	198	113	39	76	----	83	37	130	276	330	282	272	242	227	140	46	204	131	106	100	104	285	106	115	118	53	97	53
13	Little Piney Creek - SR63	2	501	472	568	549	644	1424	1423	1704	1733	1807	1823	1846	----	63	91	209	249	202	189	167	151	101	89	124	59	45	33	131	203	42	59	61	39	73	106
14	Big Piney River - SR63	2	597	579	678	661	753	1534	1532	1814	1843	1917	1933	1955	163	----	93	270	311	265	249	230	214	145	74	187	120	100	90	133	264	98	112	114	25	61	86
15	Niangua River - Bennet Springs	14	605	584	697	678	772	1551	1550	1833	1862	1934	1952	1975	486	594	----	273	273	236	208	220	202	190	161	170	135	133	121	215	224	129	147	148	83	33	176
16	Apple Creek, IL	10	278	248	346	449	409	1202	1201	1482	1511	1585	1601	1624	435	544	563	----	127	93	130	58	75	144	249	107	151	171	180	235	124	172	160	158	248	272	282
17	Middle Fabius River	1	412	381	480	583	542	1336	1334	1616	1645	1719	1735	1757	568	729	697	277	----	51	65	105	111	232	320	127	202	229	232	327	49	228	226	223	286	287	336
18	Salt River - main channel	9	380	349	447	549	510	1302	1300	1584	1611	1687	1703	1725	536	697	663	245	169	----	42	56	60	182	270	78	152	180	183	276	32	178	175	173	240	246	286
19	Middle Fork Salt River	6	451	422	520	621	583	1374	1373	1656	1683	1757	1775	1798	608	769	735	317	241	72	----	81	76	191	266	76	148	176	176	281	17	173	175	173	225	223	283
20	Sulfur Creek	1	315	286	518	486	447	1239	1238	1521	1548	1622	1640	1662	473	634	600	182	272	240	333	----	18	129	224	51	111	136	142	224	81	136	129	127	206	224	238
21	West Fork Cuivre River	1	322	291	385	492	454	1246	1244	1526	1555	1629	1645	1667	478	639	607	187	277	245	338	68	----	122	211	32	96	122	127	216	79	122	117	114	190	206	227
22	Mineral Fork Big River	1	270	240	389	441	414	1194	1193	1474	1503	1577	1593	1616	531	641	660	309	443	410	481	346	352	----	107	116	64	56	69	95	198	61	43	42	129	173	118
23	Current River - Owls Bend	10	1329	1368	1595	1362	1334	1075	1073	140	55	19	35	58	1846	1897	1915	1566	1699	1667	1738	1603	1609	1555	----	193	118	91	90	59	278	93	95	98	78	131	17
24	Loutre River	1	348	317	415	394	491	1270	1268	1551	1579	1653	1670	1693	229	336	356	282	415	381	454	319	325	377	1633	----	75	102	104	205	86	100	100	97	162	174	209
25	Bourbeuse River	1	370	340	325	541	513	1294	1292	1574	1603	1677	1693	1716	631	740	760	409	542	510	581	446	452	288	1658	476	----	28	31	137	160	26	29	27	97	126	135
26	Meramec River - Steelville Bridge	9	338	307	293	507	481	1260	1259	1542	1571	1645	1661	1683	599	708	726	377	510	476	549	414	420	256	1625	444	275	----	13	112	187	5	14	15	79	118	107
27	Meramec River - site 2	1	367	336	322	538	510	1289	1287	1571	1600	1674	1690	1712	628	737	756	406	539	507	579	444	449	285	1654	475	304	29	----	118	188	9	27	28	68	105	107
28	St. Francis River - Sam A. Baker SP	2	935	845	1201	967	941	13	11	1056	1085	1159	1175	1197	1394	1503	1522	1172	1305	1273	1345	1210	1215	1164	1139	1241	1263	1231	1260	----	290	116	108	110	133	186	54
29	Deer Creek - Fowler's Ranch	8	451	420	518	621	583	1374	1373	1656	1683	1757	1775	1798	608	769	735	317	240	72	60	311	317	481	1738	454	581	549	578	1344	----	185	185	183	239	239	294
30	Meramec River - Indian Springs	2	348	319	304	518	492	1271	1270	1553	1580	1654	1672	1695	610	718	737	386	520	488	560	425	430	267	1635	455	285	11	19	1242	560	----	18	19	77	114	109
31	Meramec River - Huzzah Valley	2	311	282	267	481	455	1234	1233	1516	1543	1617	1635	1658	575	681	700	349	483	451	523	388	393	230	1598	418	248	26	55	1205	523	37	----	3	92	132	110
32	Meramec River - Huzzah site 2	2	309	280	267	480	452	1233	1231	1514	1542	1616	1633	1654	570	679	698	348	481	449	521	386	391	229	1596	417	246	29	58	1202	521	39	----	2	94	133	113
33	Big Piney River - BS	2	573	533	642	621	718	1497	1495	1778	1806	1880	1897	1920	126	35	558	509	642	608	681	546	552	604	1860	301	703	671	700	1468	681	682	645	----	54	93
34	Osage Fork River	3	637	595	705	689	781	1559	1558	1841	1870	1944	1960	1983	188	272	621	571	705	673	745	610	615	668	1925	364	768	734	764	1530	744	745	708	707	237	----	145
35	Mill Creek - SR-M	1	1300	1337	1566	1333	1305	1046	1044	113	26	50	66	87	1817	1867	1888	1537	1669	1637	1709	1574	1580	1529	29	1605	1629	1595	1625	1109	1709	1605	1569	1568	1831	1894	----

APPENDIX 2

Population IDs, population names, drainages, sample sizes, population haplotypes (absolute frequency), and geographic coordinates for the samples examine in this study. The drainages column refers to the pattern of sequential tributaries, starting with the site of collection, which must be traversed to reach the Mississippi River. All sample sites are from Missouri, except Apple Creek, which was sampled from Illinois (IL).

ID	Population name	Drainages	*N*	Haplotype (frequency)	Coordinates	
					N Latitude	W Longitude
A	Cinque Hommes Creek	Mississippi	24	1(23), 11(1)	37.710524	89.831268
B	River Aux Vases Creek	Mississippi	30	1(30)	37.820944	90.021905
C	Harris Branch of Big River	Big-Meramec-Mississippi	21	11(19), 12(1), 17(1)	37.882534	90.503168
D	Cape Creek	Castor-Mississippi	35	35(35)	37.527957	90.166880
E	Castor River, main	Castor-Mississippi	15	35(15)	37.340264	90.209999
F	Bounds Creek	St. Francis-Mississippi	24	35(1), 36(23)	37.163321	90.449097
G	Frazier Creek	St. Francis-Mississippi	29	35(5), 36(21), 37(1), 38(1), 39(1)	37.173465	90.465033
H	Logan Creek	Black-White-Arkansas-Mississippi	31	33(24), 34(7)	36.623714	90.709938
I	Cave Spring Creek	Current-Black-White-Arkansas-Mississippi	12	22(8), 24(1), 25(2), 27(1)	36.888428	90.908867
J	Jacks Fork River at SR19	Current-Black-White-Arkansas-Mississippi	29	22(27), 23(1), 25(1)	37.154740	91.360353
K	Current River at Round Spring	Black-White-Arkansas-Mississippi	29	22(14), 25(13), 28(1), 30(1)	37.287459	91.411208
L	Jacks Fork River at SR17	Current-Black-White-Arkansas-Mississippi	22	22(12), 23(1), 25(1), 26(1), 27(1), 28(3), 29(1), 31(1), 32(1)	37.056348	91.668047
M	Little Piney Creek at SR63	Gasconade-Missouri-Mississippi	2	22(2)	37.790715	91.828168
N	Big Piney Creek at SR63	Gasconade-Missouri-Mississippi	2	22(2)	37.241847	92.009643
O	Niangua River at Bennett Springs	Osage-Missouri-Mississippi	14	9(13), 10(1)	37.743227	92.860171
P	Apple Creek (IL)	Illinois-Mississippi	10	1(10)	39.359477	90.495629
Q	Middle Fabius River	Fabius-Mississippi	1	2(1)	40.028610	91.705706
R	Salt River, main	Mississippi	9	1(9)	39.593016	91.533204
S	Middle Fork Salt River	Mississippi	6	1(5), 3(1)	39.486513	92.003082
T	Sulfur Creek	East Fork Cuivre-Cuivre-Mississippi	1	1(1)	39.194366	91.136499
U	West Fork Cuivre River	Cuivre-Mississippi	1	8(1)	39.082146	91.294877
V	Mineral Fork of Big River	Big-Meramec-Mississippi	1	11(1)	38.076182	90.739045
W	Current River at Owls Bend	Black-White-Arkansas-Mississippi	10	22(6), 25(2), 28(2)	37.181335	91.175518
X	Loutre River	Missouri-Mississippi	1	2(1)	38.888307	91.574683
Y	Bourbeuse River	Meramec-Mississippi	1	16(1)	38.223951	91.447610
Z	Meramec River at Steelville Bridge	Mississippi	9	13(7), 14(1), 15(1)	37.981501	91.369761
AA	Meramec River, site 2	Mississippi	1	13(1)	37.950500	91.510439
AB	St. Francis River at Sam A. Baker St. Pk.	Mississippi	2	18(2)	37.240440	90.511291
AC	Deer Creek	Salt-Mississippi	8	4(1), 5(5), 6(1), 7(1)	39.615534	91.900070
AD	Meramec River at Indian Springs	Mississippi	2	13(2)	37.992667	91.424840
AE	Meramec River at Hazzah Valley	Mississippi	2	13(2)	38.033359	91.224526
AF	Meramec River Hazzah, site 2	Mississippi	2	13(2)	38.057949	91.223575
AG	Big Piney River at Boiling Spring	Gasconade-Missouri-Mississippi	2	22(2)	37.462817	91.986466
AH	Osage Fork River	Gasconade-Missouri-Mississippi	3	19(1), 20(1), 21(1)	37.536009	92.589355
AI	Mill Creek at SR M	Current-Black-White-Arkansas-Mississippi	1	22(1)	37.049041	91.071359