Population genetics of the fringed darter, *Etheostoma crossopterum*, in Illinois as determined from analysis of three microsatellite loci in five populations

Final report submitted to Illinois Department of Natural Resources
Project funded through the Illinois Wildlife Preservation Fund
Grant No. 04-029W

William J. Poly
Department of Zoology
Southern Illinois University
Carbondale, Illinois 62901-6501
E-mail: argulus5@siu.edu

and

James W. Fetzner, Jr.
Section of Invertebrate Zoology
Carnegie Museum of Natural History
4400 Forbes Avenue
Pittsburgh, Pennsylvania 15213-4080

22 June 2005

Introduction

In 2001, two populations of fringed darter (*Etheostoma crossopterum*) were established in two different streams within the Cache River drainage in southern Illinois (Poly, 2003). Sixty darters (20 male, 40 female) were relocated to a single site on each of two receiving streams (Bradshaw Creek and an unnamed tributary of Cache River) on the same day of collection. Each introduction site received 13 males and 27 females from Mill Creek and 7 males and 13 females from a tributary of Big Creek. Fringed darters reproduced in both 2001 and 2002 and juveniles darters were recruited into the population in both years (Poly, 2003). Reproduction and dispersal also were monitored in 2003 and 2004. In 2002, fin clips from darters at each introduction site and tissues from other fishes originating from the source streams and other streams within the Cache River drainage were collected to compare genetic diversity among source and introduced populations and other populations of fringed darters in Illinois.

Materials and Methods

Fish communnity and darter population structure

On 20 September 2004, a seining survey was performed to determine fish community composition as well as abundance and size classes of fringed darters present at the two introduction sites by seining a 150 m section of each stream using a 3.0 x 1.8 m seine. All adult and young-of-the-year darters were measured and weighed. Total length (mm) of each darter was measured with a stainless steel ruler, and individual weights were measured to 0.1 g with a portable electronic balance. Water temperature was measured to 0.1 °C with a Taylor® digital thermometer. All captured fishes were held in coolers until a section of stream had been sampled (to avoid recaptures), then were identified, enumerated, and returned to the stream. Bradshaw Creek is a fourth order stream with an average width of 8.5 m (6.8-10.2, n = 5) within the study section (upstream of Winghill Rd., 9 km E Cobden, Union Co., Illinois (T11S, R1E, Sec. 30, 37°31′49″/89°08′58″)). Tributary of Cache River is third order and has an average width of 4.6 m

(2.1-6.2, n = 5) within the study section (at U.S. Rt. 51 bridge, 3.5 km SSE Cobden, Union Co., Illinois (T12S, R1W, Sec. 5, 37°30′12″/89°13′54″)).

Samples and DNA Extractions

In Spring 2002, fin clips were collected from 30 adult fringed darters at each introduction site from the first generation of offspring produced at the introduction sites. Fin clips were fixed and stored in 100% ethanol until extraction of DNA. In total, 192 fringed darters were sampled from 5 sites for this project (Table 1, Figure 1). Specimens examined include one unmodified population plus the source and transplant populations examined by Poly (2003). Tissue samples (fin clips and muscle tissue) were taken and shipped to the lab either in 95% ethanol or frozen on dry ice, respectively. After arriving in the lab, DNA extractions were performed using a modified high-salt precipitation method previously described (Fetzner and Crandall, 2003). In brief, the samples sent under separate conditions were initially treaded differently. Frozen tissues were placed into 1.5 ml tubes containing 600 µl of cell lysis buffer (10mM Tris-base, 100mM EDTA, 2% SDS, pH 8.0) along with 6 µl of Proteinase K (20mg/ml). For ethanol preserved tissues, samples were first transferred to a 1.5 ml tube and dried in a vacuum centrifuge to get rid of the remaining ethanol. Once dry, these samples were added to cell lysis buffer and Proteinase K identical to frozen tissues. All samples were then allowed to incubate overnight at 58°C in a heat block. The extraction then continued the following day after cooling the samples to room temperature and adding 4 µl of RNase A (20mg/ml). The samples were then incubated at 37°C for one hour, after which 300 µl of 7.5 M ammonium acetate was added and the samples were then mixed and placed on ice for 15 minutes. After 15 minutes, the samples were centrifuged at high speed for 5 minutes 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, 30 to 50 µl of TLE buffer (10mM Tris, 0.1 mM EDTA, pH 8.0) was added to resuspend the DNA. Since the

Table 1. Sampled populations of the fringed darter examined in this study. Number of specimens sampled and actual number where genotype data were obtained (N) are also listed.

No.	Population	Sampled	N	PopCode	State	County	Lat	Lon
1	Big Creek	17	17	BC	IL	Union	37.41141	-89.16424
2	Bradshaw Creek (transfer)	30	17	Brad	IL	Union	37.53028	-89.14944
3	Trib. to Cache River (transfer)	32	31	CR	IL	Union	37.50306	-89.23171
4	Mill Creek (source)	62	60	MC	IL	Union	37.36722	-89.24861
5	Trib. to Big Creek (source)	51	47	TBC	IL	Union	37.36472	-89.16333
	TOTALS	**192**	**172**					

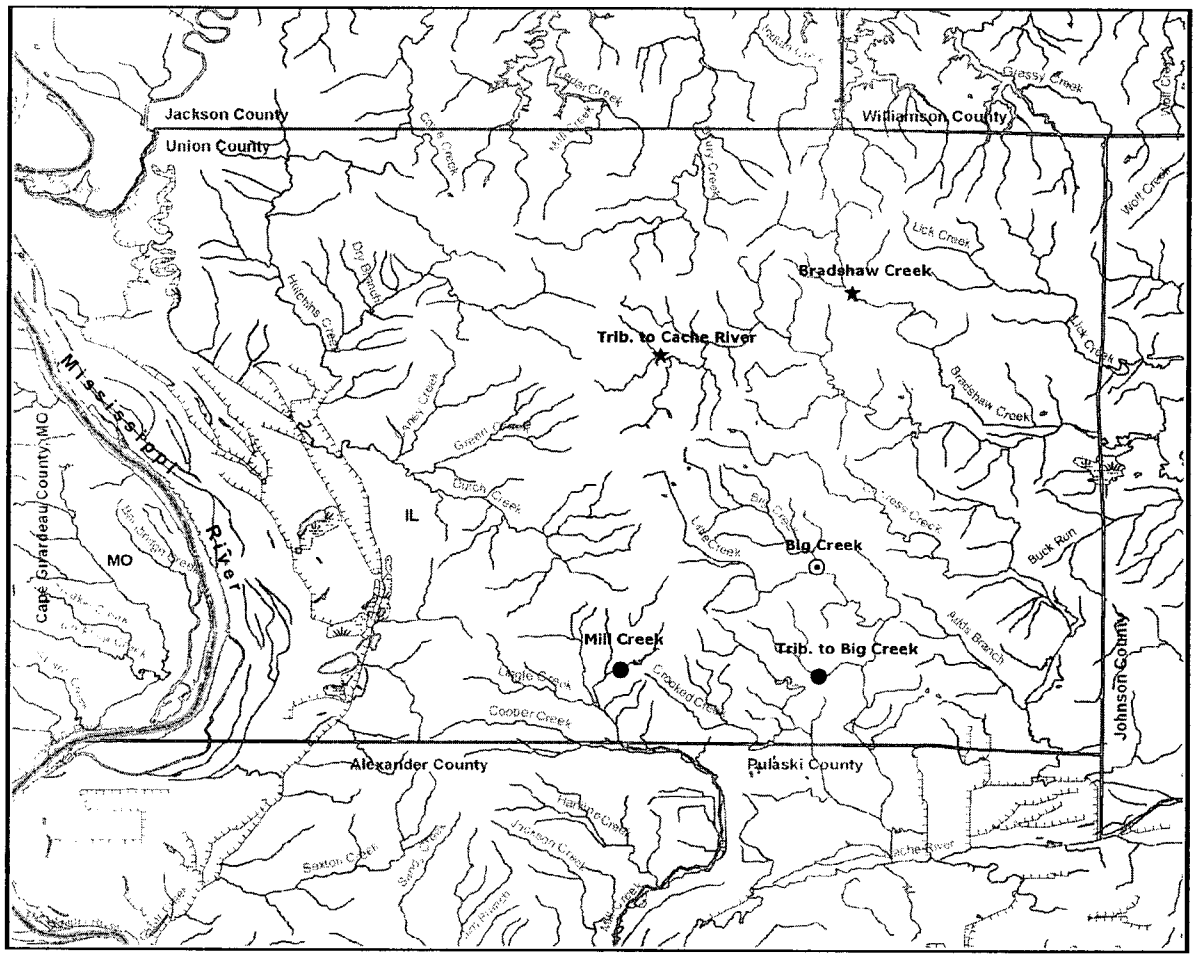

Figure 1. Map depicting collection localities for the populations examined in this study. Symbols: ⊙ = native population, ● = source populations (also native), ★ = transplanted populations.

DNA appeared to be at low concentrations and degraded in most cases, these samples were used directly in PCR reactions without dilution.

PCR and Fragment Analysis

Three microsatellite loci developed previously in the striped darter (*Etheostoma virgatum*) by Porter et al. (2002) were used to examine levels of genetic variation within and among populations of the fringed dater. Primer sequences used to amplify these loci are as follows: *CV09* F 5_-TTTCTGGTCAAGCCTCTGAG-3_, R 5_-ACAGGTGGAAGGGTCACATG-3_; *CV12* F 5_-ACTGTTAGCCCTACACTCTG-3_, R 5_-TTGAAGCAGG-TATTCTCACC-3_; and *CV24* F 5_-CTTTTGACATTGGGTTG-CATC-3_, R 5_-TCACATAGTGGGTAATGCAC-3_. Forward primers were fluorescently labeled so that PCR products could be visualized on a Beckman-Coulter CEQ2000XL capillary DNA sequencer.

Separate PCR reactions were conducted for each locus in a total volume of 10 µl and contained 1 µl each of 10x buffer, $MgCl_2$, and dNTP mix (1.25µM each) along with 0.4 µl each primer (10µM concentration; one with fluorescent label) and 0.06 µl *Taq* DNA polymerase. PCR conditions consisted of 95°C for 2 minutes followed by 35 cycles of 95°C, 15 sec, 56°C 20 sec, 72°C 20 sec followed by a hold at 72°C for 30 minutes to aid the incorporation of adenine residues to the ends of all PCR products which reduces the number of stutter bands visualized. Samples were then held at 4°C until analysis.

After PCR amplification, 50 µl (*CV12*) or 90 µl (*CV09*, *CV24*) of TLE buffer was added to each reaction, mixed and an aliquot of 10, 4, and 2 µl, respectively was added to the same well of a 96 well plate for each individual. Samples were then dried down in a vacuum centrifuge and 35 µl of deionized formamide and 0.33 µl of 400 bp fluorescently labeled size standard (Beckman-Coulter) was added to each well. All samples were then overlaid with 1 drop of mineral oil before being run on the sequencer.

Data Analysis

Genotypes at the three loci were determined using the fragment analysis software that came with the DNA sequencer. Allele sizes were calculated based on mobilities relative to a DNA size standard and then exported to Microsoft Excel for analysis. The program GenAlEx v 5.1 (Peakall and Smouse, 2001) was used to generate allele frequencies and to perform other statistical analyses such as an analysis of molecular variance (AMOVA), principal components (PCA) and distance matrix correlations (Mantel test). Other genetic analyses were performed using the Arlequin v2.001 (Schneider et al., 2000) and FSTAT v2.9.3 (Goudet, 2001) packages. Pairwise population distances were generated using the R_{ST} measure of Slatkin (1995). R_{ST} is an estimator of gene differentiation accounting for variance in allele size and is defined for genetic markers undergoing a stepwise mutation model (i.e., microsatellites). These distances were then used to perform PCA and cluster analyses.

Results and Discussion

A relatively high number of darters was collected during the Fall 2004 season, and from the length vs. weight data there appear to be at least three age classes represented as observed in natural populations of this species (Figure 2). Because the fringed darter has an extended spawning season, there is more size variation within year classes, and this variation can obscure the boundaries between year classes or result in what might appear to be two distinct year classes from a single spawning season. The fish communities of both introductions sites were similar in terms of species composition and abundance as with previous years. Fringed darter comprised 18% and 42% of the fish communities at tributary of Cache River and Bradshaw Creek, respectively, which represents an increase from 2003 but is more similar to values from 2001 and 2002 (Table 2). Eighty-six fringed darters were captured in Bradshaw Creek, whereas 88 fringed darters were collected in tributary of Cache River (Figure 2).

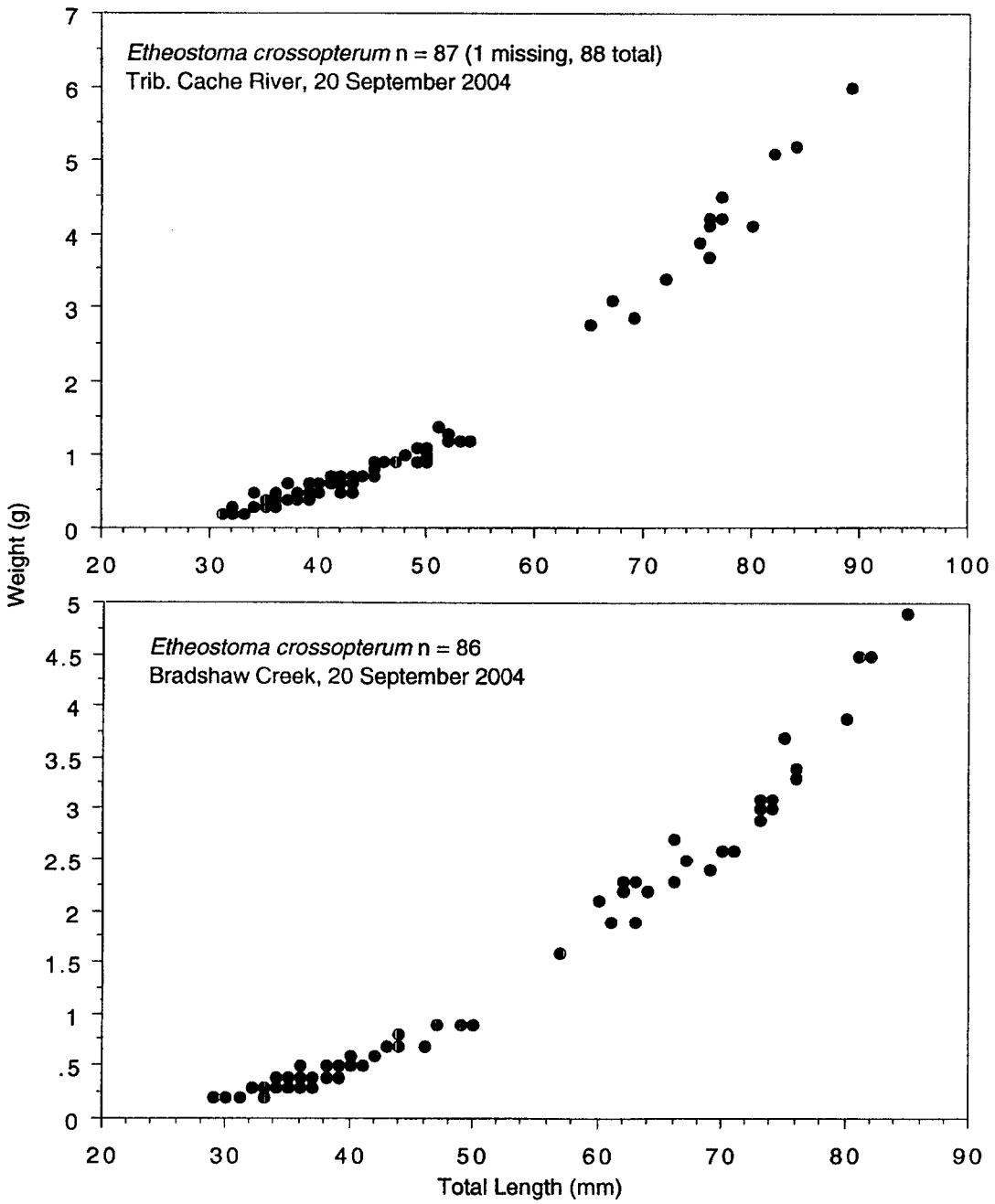

Figure 2. Weight vs length scatterplots for fringed darters in Bradshaw Creek and Tributary of Cache River in Fall 2004. Note that some data points overlap completely.

Table 2. Species composition and abundance at translocation sites (Bradshaw Creek and tributary of Cache River) in Fall 2004 (absolute abundance followed by relative abundance in parentheses).

Species	Bradshaw Creek	Tributary of Cache River
Cyprinidae		
Campostoma anomalum pullum	4 (2%)	32 (6%)
Cyprinella lutrensis	— (0%)	35 (7%)
Lythrurus umbratilis	3 (1%)	26 (5%)
Pimephales notatus	49 (24%)	120 (24%)
Semotilus atromaculatus	10 (5%)	13 (3%)
Catostomidae		
Catostomus commersoni	— (0%)	4 (<1%)
Ictaluridae		
Ameiurus melas	1 (<1%)	— (0%)
Ameiurus natalis	4 (2%)	1 (<1%)
Fundulidae		
Fundulus olivaceus	10 (5%)	46 (9%)
Aphredoderidae		
Aphredoderus sayanus	3 (1%)	7 (1%)
Centrarchidae		
Lepomis cyanellus	7 (3%)	9 (2%)
Lepomis macrochirus	5 (2%)	60 (12%)
Lepomis megalotis	12 (6%)	33 (7%)
Lepomis spp. juveniles	45 (22%)	— (0%)
Lepomis cyanellus × *L. macrochirus*	1 (<1%)	— (0%)
Micropterus salmoides	— (0%)	5 (1%)
Percidae		
Etheostoma chlorosomum	3 (1%)	— (0%)
Etheostoma crossopterum	86 (42%)	88 (18%)
Percina maculata	6 (3%)	23 (5%)
Total individuals	205	502
Total species	14 + 1	14 + 1

DNA Integrity and Amplification

Most of the DNA samples examined for this study were low in concentration and appeared to be highly degraded. As a result, some specimens completely failed to amplify via PCR, and for others the amplification of fragments from the larger microsatellite loci (*CV09*, *CV12*) was hindered. However, at least some data was recovered for all loci in all examined populations, which allowed for comparisons to be made. The Bradshaw Creek (Brad) and Tributary to Cache River (CR) parental stock populations were the most affected by the DNA degradation and amplification problems. These samples were ethanol preserved fin clips taken from the individuals released into the two new stream reaches. To avoid trouble with future studies, it is suggested that fin clip samples be handled differently, possibly by freezing in liquid nitrogen while collecting in the field.

Microsatellite Variation and Allele Frequencies

The alleles seen at the three loci examined in this study produced a variety of stutter bands when run on the DNA sequencer. However, detection of the real allelic peak was straightforward. Allelic size ranges for the three loci were 110-114 for *CV24*, 150-166 for *CV09*, and 196-230 for *CV12*. The *CV09* and *CV24* loci had repeat lengths of 2 base pairs (bp). The repeat length for *CV12* appeared to be 4 bp, but there appeared to be some other mutational processes working at this locus since alleles 226 and 230 are only 2 bp off compared to the other alleles detected at this locus. DNA sequencing of these different alleles would help to determine the cause of this shift in allele size, but was outside the scope of this study.

Table 3 shows the frequency of alleles detected in each population. In total, the *CV09* locus was found to have 2 different alleles in the samples examined with one allele (166) having the highest frequency in all populations (Figure 3A). The *CV12* locus was the most variable of the three loci examined, at least in terms of the number of different alleles detected (N=8) (Figure 3B). The most common (highest frequency) allele for this locus was not the one same for all populations (Figure 3B). The *CV24* locus contained 3 alleles with one allele (114) at high frequency in all populations (Figure 3C). Allele counts at all three of these loci were much lower than variation detected in the

Table 3. Allele frequencies detected at the three microsatellite loci examined in this study. Number of individuals (N) examined for each locus is also indicated.

	Microsatellite Loci / Alleles															
	CV09			CV12									CV24			
Population	N	150	166	N	196	200	204	220	224	226	228	230	N	110	112	114
Big Creek	17		1.000	12	0.500	0.042				0.333		0.125	17			1.000
Bradshaw Creek	12	0.167	0.833	5		0.900	0.100						17	0.059	0.235	0.706
Trib. To Cache River	29	0.052	0.948	5	0.700	0.200			0.100				31	0.096	0.016	0.887
Mill Creek	50	0.230	0.770	38	0.079	0.816	0.105						60	0.367	0.017	0.616
Trib. To Big Creek	43		1.000	31	0.533	0.065		0.161	0.080		0.161		47		0.032	0.968

Table 4. Estimation of overall heterozygosity values according to Nei (1987). H_O is the observed proportion of heterozygotes; H_S is the within sample gene diversity; H_T is the overall gene diversity; H_T' is H_T, but independent of sample size; D_{ST} is the amount of gene diversity among samples; D_{ST}' is D_{ST} but independent of sample size; G_{ST} is an estimator of F_{ST} (population subdivision); G_{ST}' is G_{ST} but independent of sample size; and G_{IS} is an estimator of F_{IS} (amount of inbreeding within subpopulations).

	Diversity Indices								
Locus	H_O	H_S	H_T	H_T'	D_{ST}	D_{ST}'	G_{ST}	G_{ST}'	G_{IS}
CV09	0.163	0.149	0.164	0.168	0.015	0.019	0.091	0.111	-0.097
CV12	0.390	0.477	0.701	0.757	0.224	0.280	0.320	0.370	0.182
CV24	0.265	0.243	0.288	0.300	0.045	0.057	0.158	0.190	-0.093
Overall	0.273	0.290	0.384	0.408	0.095	0.119	0.247	0.290	0.058

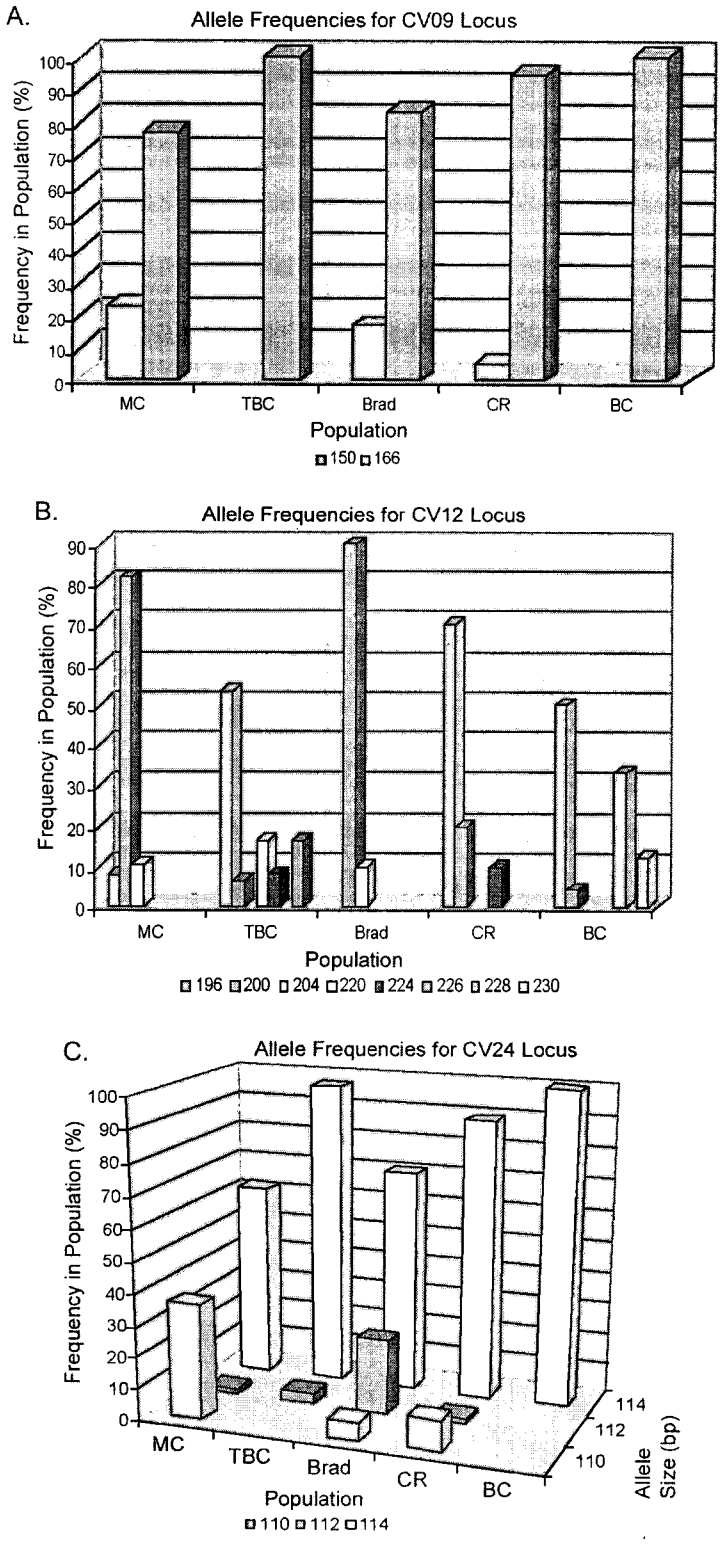

Figure 3. Allele frequency distributions for all populations for A. *CV09* locus, B. *CV12* locus, C. *CV24* locus.

striped darter (*Etheostoma virgatum*), the species from which these loci were originally isolated (Porter et al., 2002). In comparison, allelic counts among only two populations of the striped darter for *CV09*, *CV12*, and *CV24* were 22, 24, and 36 alleles, respectively (Porter et al., 2002).

Gene diversity within populations for each locus ranged from a low of 0.000 (no variation) to a high of 0.668. Estimates of overall genetic diversity values are listed in Table 4.

Population Differentiation

Figure 3 suggests that the makeup of the transplanted populations are actually a bit different from one another, even though the same number of individuals were used from each source population in the relocation effort. It appears from the data for CV12, that the Brad population is most similar to MC, while CR is closer to TBC. However, since the number of individuals we were able to genotype for the Brad and CR populations was only a fraction of those original individuals actually stocked at the new localities, these results would seem to be suspect. The other two loci clearly show that MC made a contribution to both populations as can be seen by the transfer of the unique MC alleles (*CV09*:150 allele; *CV24*:110 allele) which were not found in the populations from the Big Creek drainage system. Since the number of individuals selected for transfer was relatively low (60 individuals released at each new site), the possibility always exists that these two populations might show marked differences from one another in allelic frequencies mainly due to sampling error (founder effects). Parental stocks originally transplanted into Brad and CR could be checked for introgression by checking subsequent generations for mixing of alleles. Private alleles were detected in each source population (see Figure 3), which could aid in identification or assignment to specific populations.

An analysis of molecular variance (AMOVA) suggested that 82% of the genetic variation was located within individual populations while only 18% was associated with among population variation (R_{ST}=0.175, P=0.001). This value is analogous to F_{ST}, or the degree of population subdivision. Cluster analysis of pairwise R_{ST} values among populations grouped the two transplanted populations together and placed the source populations together (Figure 4). The

Mantel test for the correlation of geographic and genetic distances was not significant ($R^2 = 0.599$, P=0.080).

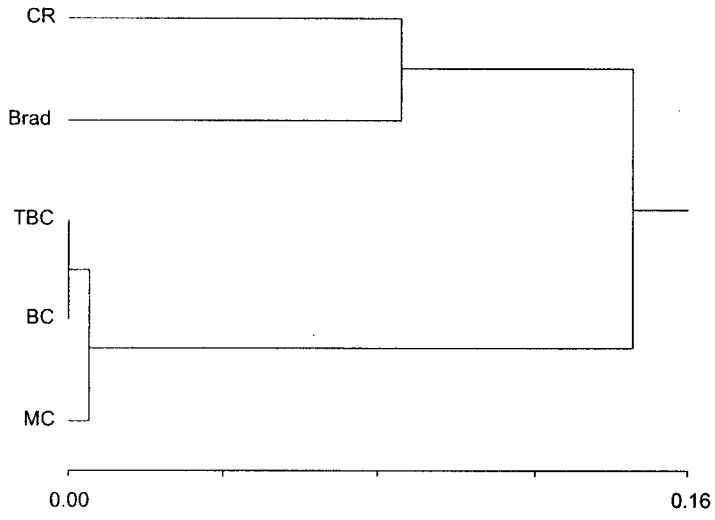

Figure 4. UPGMA phenogram of population relationships based on pairwise R_{ST} distances among populations.

Summary

In 2001, darters were collected from two streams in the Cache River basin, and a combination of darters from each source stream were moved to the receiving streams, which also were in the headwaters of the Cache River basin (20 males and 40 females to each). One year later, fin clips from 30 fish at each introduction site and tissues from other darters from the source streams and other streams within the Cache River drainage were collected. Genetic comparisons among (1) the two introduction populations, (2) the two source populations, and (3) one other population from Illinois were made to determine variability within and among populations, and in particular, to determine if the introduced populations maintained levels of heterogeneity similar to source populations or if a loss of genetic variation resulted. Some genetic differences among populations were detected even in the small area under study. Microsatellite alleles from the two source streams were present in both introduced populations, which accounts for the genetic similarity of these two

populations. However, small samples sizes prevent an accurate understanding of the relative contributions of the source populations to introduced populations at this time. In surveys made at the two introduction streams in September 2004, fringed darters were present in relatively high numbers at both sites, indicating persistent reproduction and maintenance of healthy populations over a four-year period.

Acknowledgments

The authors appreciate the help of James E. Wetzel and Karen Gibson (Southern Illinois University) for assisting with fieldwork and ordering equipment and supplies, respectively. The Illinois Department of Natural Resources granted permits for the project and provided funding through the Illinois Wildlife Preservation Fund.

Literature Cited

Fetzner, Jr., J.W. and K.A. Crandall. 2003. Linear habitats and the Nested Clade Analysis: An empirical evaluation of geographic vs. river distances using an Ozark crayfish (Decapoda: Cambaridae). Evolution 57: 2101–2118.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Peakall, R. and P. E. Smouse. 2001. GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia. Available from: http://www.anu.edu.au/BoZo/GenAlEx/.

Poly, W.J. 2003. Design and evaluation of a translocation strategy for the fringed darter (*Etheostoma crossopterum*) in Illinois. Biological Conservation 113: 13–22.

Porter, B.A., A.C. Fiumura, and J.C. Avise. 2002. Egg mimicry and allopaternal care: two mate-attracting tactics by which nesting striped darter (*Etheostoma virgatum*) males enhance reproductive success. Behavioral Ecology and Sociobiology 51: 350–359.

Schneider, S., D. Roessli, and L. Excofier. 2000. Arlequin: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Department of Anthropology, University of Geneva. Available from: http://anthro.unige.ch/arlequin/.

Slatkin, M. 1995. A Measure of Population Subdivision Based on Microsatellite Allele Frequencies. Genetics 139: 457–462.

Budget for project:

Funds from Illinois Department of Natural Resources,

Illinois Wildlife Preservation Fund		estimated	actual
Genetic analyses		$700.00	$667.00
Indirect (SIUC admin. cost)		----------	$133.00
Mileage		$100.00 [for 300 mi.]	----------
	Total:	**$800.00**	**$800.00**

Funds contributed by William J. Poly		estimated	actual
Dry Ice for shipping tissue samples			$16.76
Fed Ex shipping (tissue samples)			$88.15
Travel			$320.00
Acculab Pocket Pro 250B electronic balance	qty. 1		$166.20*
Report preparation			$80.00
	Total:		**$671.11**

*equipment from an earlier study