EVALUATION OF MICROSATELITE LOCI IN LARGEMOUTH BASS (*Micropterus salmoides*): RESOLUTION OF POPULATION STRUCTURE AND INDIVIDUAL ORIGINS

THESIS

Presented to the Graduate Council of

Texas State University-San Marcos

in Partial Fulfillment of

the Requirements

For the Degree

Master of SCIENCE

By

Dijar J. Lutz-Carrillo, B.S.

San Marcos, Texas

May 2004

DEDICATION

This thesis is dedicated to my family here and passed, my wife and hers, and my friends.

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the contribution of a number of people on this research project. My co-advisors, Dr. Timothy Bonner and Dr. Chris Nice, provided me with invaluable amounts of their time and contributed a great deal of insight on this project. I would also like to thank my committee member, Dr. Mike Forstner, for his insight into molecular genetics and constructive suggestions for experimental design and analyses.

Thanks go to the following: Loraine Fries, Jamie Dixson, Greg Southard, Beverly Villarreal, Allan Forshage and the rest of the Texas Parks and Wildlife Department for providing invaluable suggestions, working around my schedule, continually finding funding for the project and for patiently bringing me along. Todd Kassler, John Epifanio, and Dave Philipp of the Illinois Natural History Survey, Bill Wentroth of the Oklahoma Department of Wildlife Conservation, and Bonnie Brown of the Virginia Commonwealth University for providing samples and microsatellite sequences. My peers and fellow graduate students, Casey, Dusty, Brad, Val, Jackie, Tracy, Michele, Julie, Karma, Brian, Lene and Kackie, thank you for asking questions and for your company on many exciting trips. My wife and my family thank you for enduring and supporting my many years of school and my many hours of studying.

The current research project was wholly funded through grants from the Parks and Wildlife Foundation of Texas and the Texas Parks and Wildlife Department.

TABLE OF CONTENTS

Page

APPENDIX A

LIST OF TABLES

APPENDIX B

LIST OF FIGURES

ABSTRACT

EVALUATION OF MICROSATELITE LOCI IN LARGEMOUTH BASS (*Micropterus salmoides*): RESOLUTION OF POPULATION STRUCTURE AND INDIVIDUAL ORIGINS

By

Dijar J. Lutz-Carrillo, B.S.

Texas State University-San Marcos

May 2004

SUPERVISING PROFESSORS: Timothy H. Bonner and Chris C. Nice

Microsatellite DNA variation was evaluated for the first time within the largemouth bass (*Micropterus salmoides*). Eleven of forty-five primer pairs, originally developed for the amplification of microsatellite loci in other centrarchid species, were found to be suitably polymorphic, or expressed conserved differences in allele frequencies between subspecies of *M. salmoides* ($H_E = 0.407$), for utilization in this study. Based on genetic variation at these loci and concurring measures of genetic distance and differentiation, multiple ANOVA analyses, and Bayesian inference techniques, population structure within the species was resolved as consisting of three genetic assemblages with *M. s. floridanus* (Florida largemouth bass) in peninsular Florida, and *M. s. salmoides* (Northern largemouth bass) at northern and southern latitudes within North America. Furthermore, samples from the Devil's River (southwest Texas) expressed unique allele frequencies at four of eleven loci. This fine scale genetic resolution has not previously been reported within the species and, given previous reports of morphologically distinct characteristics for the Devil's River population, suggests the unique nature of this group within *M. salmoides*. Additionally, a new method for probabilistically assessing the subspecific, regional and in some cases the population status, and genomic composition, of individuals within the species, allowing the correct assignment of 99% to 100% of individuals is provided. Although subspecies were confidently resolved, high levels of introgression would greatly complicate the process of identification. Simulation studies indicated that the clustering values of intergrades approach those of pure subspecies after two parental backcrosses from an F_1 generation. Microsatellite loci were also assessed for their utility in assigning multilocus fingerprints to individuals. Among 473 individuals, no duplicate genotypes were identified, but β-

error probabilities ranged from 2.8×10^{-3} to 2.7×10^{-7} within populations. These relatively high levels of duplication and large number of loci required would limit the practicality of this application, but levels of polymorphism at two loci, Lma021 and Lar007, suggest that the future identification of other highly polymorphic loci will make fingerprinting and thus the unique identification of individuals feasible.

INTRODUCTION

As recreational stocks have proven their revenue potential, the management of fish species has intensified, leading to increasing anthropogenic influences on the distribution and density of many fishes (Allendorf 1991; Bianchini 1995; Shafland 1999). The largemouth bass *Micropterus salmoides* typifies this and has become one of the most intensely managed and pursued sport fish in the United States (Beaty and Childers 1980; USFWS 1996; USFWS 2001). Two allopatric subspecies, *M. s. salmoides* and *M. s. floridanus*, sometimes considered separate species (Kassler et al. 2002), are thought to have diverged during the late Pliocene as a result of significant sea level fluctuations fragmenting their native tributaries (Near et al. 2003). The subspecies were recognized and described in 1949, with *M. s. floridanus* confined to peninsular Florida and the range of *M. s. salmoides* extending to the mid-western and eastern half of the United States, southern Quebec and Ontario, and northeastern Mexico (Bailey and Hubbs 1949). They depict the region between the subspecies as consisting of an intergrade or secondary contact zone. Due to large scale translocations, their ranges now encompass most of North America and parts of Japan, Europe, Africa and South America (MacCrimmon and Robbins 1975; Philipp et al. 1983).

Public demand for specific characteristics of the two subspecies has driven translocation efforts which have changed the distribution and composition of *M. salmoides* populations. In southern latitudes, *M. s. floridanus* exhibits faster growth rates

after the first year of development (Clugston 1964; Johnson 1975; Philipp 1991; Philipp and Whitt 1991; Smith and Wilson 1981; Zolcynski and Davies 1976), and attains a greater overall size (Bottroff and Lembeck 1978; Forshage and Fries 1995; Inman et al. 1978; Maceina and Murphy 1992; Wright and Wigtil 1981). The Florida subspecies has also been credited with increasing the state record and the number of reservoirs yielding trophy fish (>4.5 kg) in Texas (Forshage and Fries 1995), but it is significantly less susceptible to angler harvest than *M. s. salmoides* (Kleinsasser et al. 1990; Zolcynski and Davies 1976). In northern latitudes however, *M. s. salmoides* is superior in terms of longevity and growth (Fields et al. 1987; Philipp and Whitt 1991). It should also be noted that Kleinsasser et al. (1990) found that *M. s. floridanus* female x *M. s. salmoides* male intergrades surpassed both subspecies in terms of size and condition in Texas ponds.

A significant investment has been made in *M. salmoides* to produce quality fisheries (Bottroff and Lembeck 1978; Forshage and Fries 1995; Philipp 1991), yet in order to apply management strategies that account for the biological constraints of each subspecies, and balance conservation of the genetic integrity of assemblages with public demand, a better understanding of population structure and the ability to accurately identify the origins of individuals is vital. Morphological characters are unreliable and inefficient, requiring the sum of numerous meristic counts, and these difficulties are only exacerbated by introgression. Allozymes (Philipp et al. 1983; Williamson et al. 1986), restriction fragment length polymorphisms (RFLP) of mitochondrial DNA (mtDNA) (Johnson and Pignature 1995; Nedbal and Philipp 1994), and randomly amplified polymorphic DNA (RAPD) (Williams et al. 1998) have also been utilized to discriminate between the subspecies of *M. salmoides*. However, these markers remain unsuitable for

stock identification due to the number of loci identified, levels of polymorphism, the mode of parental transmission, repeatability, and the direct and indirect mortality associated with sampling techniques (Avise 1974; Harvey et al. 1984; Lewin 1989; Tibayrenc et al. 1993).

Microsatellites offer a potentially superior alternative in that they are ubiquitously distributed throughout the genome (Tautz and Renz 1984), display high levels of polymorphism (Jeffreys et al. 1994), and require only small amounts of DNA for successful screening, thus reducing the need for invasive and consumptive sampling and allowing analyses of degraded samples (Nielsen et al. 1997; Nielsen et al. 1999; Hutchinson et al. 1999). As a consequence they have become increasingly popular genetic markers for resolving population structure and identifying the origins of individuals.

Microsatellites consist of internally repetitive sequences of DNA, the most common motif being $(Cytosine/Adenine)_n$ (Arnaud et al. 1993; Wright 1992), many of which have been duplicated or translocated generously throughout the genome (Wilder and Hollocher 2001). They are presumed to be neutral (Tautz et al. 1986), co-dominant genetic markers, which appear to conform to Hardy-Weinberg Equilibrium (Colbourne et al. 1996), though they violate the model's assumptions regarding mutation rates. Their internally repetitive structure imparts a high rate of mutation, relative to nuclear genomic coding sequences, due to slipped-strand mispairing during DNA replication and unequal crossing over during meiosis (Jeffreys et al. 1994; Levinson and Gutman 1987). The resulting expansions and contractions convey high levels of variation to microsatellite loci, but may also result in the loss of the sequence and make them susceptible to size

homoplasy, which can lead to underestimates of population subdivision and genetic divergence among populations and species (Estoup et al. 1995; Garza and Freimer 1996; Taylor et al. 1999).

Despite these concerns microsatellites have been effective for studying population sub-structure in other teleosts (Landry and Bernatchez 2001; Vasemagi et al. 2001) and genetic fingerprinting in other centrarchids, allowing the discrimination of individuals and providing a high degree of resolution for parentage determination using as few as four markers (Colbourne et al. 1996, DeWoody et al. 2000). The aquaculture industry has also made use of microsatellites for an *Ictalurus punctatus* marker-assisted breeding program (Waldbieser and Wolters 1999), and Colbourne et al. (1996) utilized microsatellites in the study of fish behavior in *Lepomis macrochirus*.

Although the microsatellite sequence is highly variable, their flanking regions are relatively conserved, allowing cross-species amplification using identical primer pairs (Colbourne et al. 1996; Neff et al. 1999; Weber and May 1989). However, heterozygosity and polymorphism commonly decrease with increasing phylogenetic distance from the taxon in which the loci were first isolated (Ellegren et al. 1995). For this study microsatellite primer pairs originally developed for other centrarchid species were surveyed for variation. Loci were screened to specifically select for high levels of polymorphism and their power to distinguish subspecies of *M. salmoides*. A reduction in motif repeats, observed to occur when moving from a species of origin to other taxa, is not a concern here given the *a posteriori* selection of polymorphic loci and the assertion of Crawford et al. (1998) that ascertainment bias is only a minor contributor to variation at microsatellite loci after observations across two vertebrates (domestic cattle and sheep)

at 472 microsatellite loci. The purpose of this study was to quantify microsatellite population genetic structure geographically and taxonomically within *M. salmoides*, refine accepted management units, assess the utility of centrarchid microsatellite loci for use in resolving individual origins and assigning unique genetic fingerprints to individuals.

MATERIALS AND METHODS

Samples

Fin clips, muscle tissue or blood samples were taken from 473 fish representing fourteen *M. s. salmoides* and *M. s. floridanus* populations (Table 1, Figure 1). Samples of *M. s. floridanus* were obtained from the Hillsboro River, FL, ($N = 35$), East Lake Tohopekaliga, FL ($N = 51$), Lake Medard, FL ($N = 34$), Lake Kissimmee, FL ($N = 27$) and a Texas State Hatchery population from A. E. Wood ($N = 72$). These samples were chosen as representative samples of the Florida subspecies based on geographic location and previous studies showing the Hillsboro River, East Lake Tohopekaliga, and Lake Kissimmee populations to consist of only Florida alleles based on allozyme analysis (Kassler et al. 2002; Philipp et al. 1983). The Hillsboro Canal, an extension of the Hillsboro River, and Lake Tohopekaliga, into which East Lake Tohopekaliga flows, have also been shown to consist of only *M. s. floridanus* mtRFLP haplotypes (Bremer and Zhang 1998; Nedbal and Philipp 1994). Samples of *M. s. salmoides* were obtained from Lake Kickapoo, TX ($N = 27$), Lake Charlotte, OK ($N = 50$), Twin Oaks Reservoir, TX ($N = 31$), Lake Fryer, TX ($N = 30$), Lake Marvin, TX ($N = 5$), the Devil's River, TX ($N = 37$), Lake Minnetonka, MN ($N = 27$), Lake Pepin, MN ($N = 24$), and Pike Lake, WI ($N = 23$) (Table 1; Figure 1). These samples were chosen based on geographic location and official stocking records indicating that no *M. s. floridanus* stockings had occurred. The sample sizes used approximate the general recommendation for using 30 to 50 individuals per sample when analyzing population

structure with microsatellite data (Bernatchez and Duchesne 2000; Blouin et al. 1996; Ruzzante 1998; Sjogren and Wyoni 1994). Fin clips were stored in 70% ethanol (EtOH) where they remained at room temperature until DNA extraction.

DNA Extraction and Quantification

Genomic DNA was extracted from each fin following a modified version of the Purgene protocol for extraction from fish tissue (Gentra Systems, Inc., Minneapolis, MN, USA). Approximately 3-5 mm^3 of fin tissue was placed in 300 μl of cell lysis solution (10mM tris-HCl, 10mM EDTA, pH 8.0 and 2% SDS and 3 μl [20 mg/ml] proteinase K was added. This was incubated at 55°C for 1.5-2 hours, and then cooled to room temperature, where 120 μl of protein precipitation solution (7.5 M Ammonium Acetate) was added and well mixed. The solution was incubated at 0°C for 10-15 minutes and centrifuged at 13,000 Xgravity for 5 min. The supernatant was added to 1,000 μl of 100% EtOH, mixed and stored at -80°C for 10 min, and re-centrifuged at 13,000 Xgravity for 10 minutes. The supernatant was then decanted and the pellet washed with 600 μl of 70% EtOH; this was centrifuged at 13,000 Xgravity for 5 min, decanted, and dried at room temperature for 15 min. The DNA was re-suspended in 200 μl dH_2O and left to re-hydrate at room temperature for 24 h. For each sample 4 μl DNA were mixed with 2 μl loading dye and electrophoresed on 0.8% agarose gels in a 1X TAE buffer (40 mM tris-acetate, 1 mM EDTA). Each run was visualized under UV light with the incorporation of 2 μl Ethidum Bromide (10 mg/ml) to the gel, and quantified against calf thymus DNA at concentrations of 1.0, 0.5, 0.25, and 0.125 μg/ml (Gentra Systems, Inc.,

Minneapolis, MN, USA) or a 1Kb ladder. Aliquots were then diluted 1:32 before microsatellite analysis; all isolated DNA was stored at -30°C.

Primers and PCR Optimization

Forty-five primer pairs designed for the amplification of microsatellite loci in the centrarchids *Lepomis macrochirus* ($N = 24$) (Colbourne et al. 1996; Neff et al. 1999), *L. auritus* ($N = 2$) (Malloy et al. 2000), *Enneacanthus chaetodon* ($N = 5$), *M. dolomieu* ($N = 11$) (Malloy et al. 2000), and *M. salmoides* ($N = 3$) (DeWoody et al. 2000) were screened for their utility in amplifying microsatellite loci in *M. salmoides* subspecies. Each primer pair was subject to optimization using a series of buffers varying in stringency from 0.0075 to 0.0175 M $MgCl_2$, pH 8.5 to10.0, and a temperature range of ± 8°C of the primer's reported duplex melting temperature. A standard polymerase chain reaction (PCR) profile of 60 (94°C)/60 (T_A)/180 (72°C) seconds for 30 cycles was utilized in most cases. All PCR amplifications were performed using an MJ Research PTC-200 thermocycler (Gentra Systems, Inc., Minneapolis, MN, USA). Initial PCR products were electrophoresed on 2% agarose gels to detect any locus amplification. Primer pairs that amplified *M. salmoides* loci were then obtained with an M13(-29) adapted forward primer. In this procedure, the forward primer was synthesized with an M13(-29) nineteen base sequence at the 5′-end that acted as a template for an infrared dye (IRD) labeled M13 primer to use during PCR amplification. The reactions were then re-optimized in 10 μl reaction volumes under the conditions described above and standard reaction mixtures containing 1-3 ng of template DNA, 0.04 μM of the M13(-29) adapted forward primer, 0.2 μM of the reverse primer, 200 μM dNTP's, 1X reaction buffer, 0.25 units of

Promega's *Taq* polymerase, and 0.02 μM of 25% labeled M13(-29) IRD 700/800 sequence (Li-Cor Inc., Lincoln, NE, USA). Once PCR amplified, 5 μl of stop solution (4 parts formamide to 1 part stop dye) were added to each reaction followed by denaturation at 94°C for 3 min. One-half microliter of each sample was loaded on to a previously prepared 6.5% denaturing polyacrylamide gel and electrophoresed on a Li-Cor 4200 DNA sequencer (Li-Cor Inc., Lincoln, NE, USA) in 1X TBE (0.89 mM tris-borate, 2 mM EDTA, and pH 7.4) along side a florescent labeled 350 bp ladder. Loci indicating suitable levels of polymorphism (≥ 4 alleles/ locus) or observable differences in allele frequencies between subspecies, in sub-sample runs (n = 16), were deemed acceptable and utilized in genotyping all individuals. Each individual was genotyped at 11 loci using GeneImagIR (Scanalytics, Billerica, MA), an allele scoring program.

Analyses

Tests for Hardy-Weinberg Equilibrium at each locus were performed using exact tests with a modified Markov chain algorithm in *Arlequin* version 2.0 (Schneider et al. 2000) (10,000 dememorization, 100,000 Monte Carlo (MC) steps; Guo and Thompson 1992) and *GenePop* version 1.2 (Raymond and Rousset 1995). *Arlequin* was also used to estimate linkage disequilibrium between pairs of loci using an extension of Fisher's exact probability test (Slatkin 1994; Slatkin and Excoffier 1996; Lewontin and Kojima 1960). Statistical significance levels ($\alpha = 0.05$) were modified for multiple comparisons using sequential Bonferroni adjustments (Rice 1989). Loci indicating heterozygote deficiencies were subsequently analyzed for the presence of null alleles and, if appropriate, corrected genotypes were produced using *Micro-Checker* Version 2.2.0

(http://www.microchecker.hull.ac.uk/). This software applies excess homozygotes over each allelic class to infer the presence of null alleles. Null allele frequencies, and corrected genotypes, were then estimated according to HWE. These results were compared to estimates of Chakraborty et al. (1992) and Brookfield (1996) within *Micro-Checker* in order to identify any discrepancies in the data. Similarly, these estimates were based on the presence of heterozygote deficits and HWE and used to correct null allele frequencies and genotypes. Additionally, mean levels of heterozygosity and genetic diversity were compared with previous publications at these loci.

Genetic diversity within populations was recorded as the number of alleles per locus (A), expected heterozygosity (H_E), and observed heterozygosity (H_O) (Table 2). Measures of genetic differentiation among populations were calculated by *Arlequin* using approaches based on two modes of mutational processes. Both, pairwise F_{ST} values, based on the infinite alleles model (IAM) (Kimura and Crow 1964; Michalakis and Excoffier 1996; Weir and Cockerham 1984) and pairwise R_{ST} values, based on the step-wise mutation model (SMM) (Ohta and Kimura 1973; Slatkin 1995; Michalakis and Excoffier 1996; Rousset 1996), were calculated to evaluate results between methods, given the unsatisfactory understanding of mutational modes at microsatellite loci (Jin et al. 1996). Both methods were specifically utilized to explore levels of variation within and between subspecies.

Genetic distances were calculated to identify relative relationships among populations. Cavalli-Sforza and Edward's chord distance (1967) as calculated in *Phylip* version 3.6 (Felsenstein 1989), was utilized based on the model's assumption of no underlying mutational method. The resultant matrix of pairwise distances was then

displayed as a UPGMA tree and using non-metric multidimensional scaling techniques (NCSS), which computes coordinates for each population so that each point fits proportionally to the measured distance (Lessa 1990). This allows exploration of within-species, non-hierarchical, geographic patterns, such as clinal variation and introgression. Furthermore, multi-dimensional scaling techniques do not assume linearity as other coordinate techniques, like principle component analysis, and thus may be more appropriate for analyzing genetic distances since allele frequencies at each locus have an upper limit of 1. Lessa (1990) cautioned that the use of linear models with non-linear variables may distort resultant relationships among populations. Cavalli-Sforza and Edward's distances were also employed to construct a UPGMA tree in *Phylip* illustrating the relationships among populations.

Hierarchical analyses of molecular variance were carried out using *Arlequin* to partition molecular and allele frequency variance at various geographic and genetic scales (Excoffier et al. 1992), for example *a priori* groups based on geographic location and alluded subspecies status. Groups were also defined based on clades depicted from estimates genetic distance and differentiation to form regional groups and less inclusive levels of organization. In the first analysis of variance (ANOVA), populations were grouped according to *a priori* knowledge of subspecies status. In the second ANOVA, populations were grouped into three regional groups based on estimates of genetic distance and differentiation. These groups consisted of the *M. s. floridanus* populations and *M. s. salmoides* populations located at northern and southern latitudes. In the third ANOVA, individuals were partitioned into those sampled from the Devil's River and those individuals sampled from the remaining *M. s. salmoides* localities, based on allele

frequency differences identified with estimates of genetic distance and differentiation. Additional ANOVA analyses were performed at consecutively less inclusive levels of organization to validate the results of the model. These analyses were first performed for all loci and then on a locus-by-locus basis in order to assess the differential impact of specific loci on the observed patterns. This allowed the exploration of both the structure of the data set and the loci most responsible for the observed divisions.

The assignment of individuals to multiple hierarchical levels of most probable origin was conducted using the Bayesian method implemented in *Structure* version 2 (Pritchard et al. 2000) and the log-likelihood method implemented in *Arlequin* (Paetkau et al. 1995; Paetkau et al 1997; Waser and Strobeck 1998). The assignment tests used here are similar but not identical. The Bayesian approach creates clusters of related individuals based on their multilocus genotypes, while simultaneously estimating population allele frequencies, and thus there is no tautological bias in the results. In this case there is a desire to determine whether the structure of genetic variation within *M. salmoides* is congruent with current taxonomy. The classical log-likelihood assignment test implemented in *Arlequin* estimates allele frequencies assuming the existence of distinct populations and uses those parameters to infer the source populations of unknown individuals (Buchanan et al. 1994; Paetkau et al. 1995; Rannala and Mountain 1997; Cornuet et al. 1999; Pritchard et al. 2000).

Within *Structure* multiple models were utilized, evaluating all possible clusters under Hardy-Weinberg Equilibrium, to identify clustering solutions of the highest likelihood (Rosenberg et al. 2001), allowing not only the identification of population structure, but the composition of each cluster on an individual basis, and the clustering

probability (i.e. genomic composition) of each individual. For specific analyses, for instance subspecies analysis, the number of clusters was assumed (i.e. k = 2 in this case), best representing *a priori* knowledge of genetic structure (i.e. subspeciation) in *M. salmoides* (Bailey and Hubbs 1949; Kassler et al. 2002; Near et al. 2003). Such analyses were used in order to corroborate the subspecific status of our samples. The number of incorrect assignments (i.e. *M. s. floridanus* < 0.5 or *M. s. salmoides* > 0.5) was based on assumed *a priori* knowledge of the population's subspecific status given the geographic location, stocking records, and previous studies of the samples, and was extrapolated to estimate the probability of incorrect assignments (P_I) using this method of subspecies identification. The power of assignment success was also evaluated at two levels of stringency, $P_{R\,0.95}$ and $P_{R\,0.99}$, which represent the proportion of individuals rejected from their cluster given a clustering probability ≤ 0.95 and ≤ 0.99 respectively. To quantify increases in discriminatory power four to eleven loci were employed in order of discriminatory power based on locus-by-locus ANOVA results at the hierarchical level of subspecies.

To cross-validate subspecies identification, the model constructed in the analysis above, with assignment probabilities at k = 2, was used to assign each of the individuals in the Lake Kissimmee and Lake Pepin populations. These samples were omitted from the original model and added using the step-wise model described by Efron (1983) and used by Easer and Stroebeck (1998). The proportions of incorrect assignments and the proportion rejected from their cluster at each stringency level in these two groups were then extrapolated to estimate the P_I, $P_{R\,0.95}$, and $P_{R\,0.99}$ of unknowns. The assignment of individuals to subspecies origin was then repeated in *Arlequin* by calculating the log-

likelihood of each individual's genotype arising in either group. The P_I was calculated for this method as well and used to corroborate results and identify any incongruities.

Given the degree of natural and induced introgression recorded between the subspecies (Philipp et al. 1983), it was also of interest to determine if intergrades could be resolved from pure subspecies. For this, generations of artificial intergrades were simulated between pairs of potential parental populations of opposite subspecies using the program *Hybridlab* version 0.9 (Nielsen et al. 2001). The program produces multilocus intergrade genotypes between two populations, drawing alleles at random. Simulated pairwise crosses of Hillsboro River samples with Lake Minnetonka, the Devil's River, and Lake Fryer, and Lake Kissimmee samples with Lake Minnetonka, the Devil's River, and Lake Fryer, were used to create 180 F_1 eleven locus genotypes. These populations were chosen as best representing variation within subspecies and among regions based on results from analyses of genetic distance and subdivision. These simulated F_1 genotypes were then backcrossed to each of their respective parental populations for three generations to simulate the most probable impact of a stocking event, where fewer fish are stocked than inhabit the system and multiple parental backcrosses may occur. All simulated individuals were then assigned to potential source populations using *Structure* based on their clustering scores under the subspecies model of $k = 2$, with admixture, and plotted on a scale from 0 to 1. The distribution of clustering scores for each generation of simulated intergrades was then z-score normalized to allow the accurate assessment of introgression levels (subspecies composition) in individuals. Overlapping distributions of clustering scores, based on the

mean ± 3 σ of each group (99.7% of normalized individuals), were used to quantify the error rate when using this marker suite in introgressed populations.

Bayesian inference techniques implemented in *Structure* were also used to assign individuals to less inclusive assemblages identified within *M. salmoides*. Assignment success was quantified according to the clusters expected based on measures of genetic distances and differentiation among populations, and estimates of P_I, $P_{R\ 0.95}$, and $P_{R\ 0.99}$ were obtained as described previously for subspecies assignments. Individuals were also assigned likelihood scores for each assemblage based on the log-likelihood assignment test implemented in *Arlequin*. These results were compared to the expected divisions and estimates of P_I were made as well, to corroborate the results of Bayesian analysis. Contradictory results among the analyses were recorded as indications of weak regional divisions, whereas concurring results were indicative of strong divisions.

To assess the utility of this marker system for differentiating (fingerprinting) individuals within populations, regions, subspecies, and the species, allele frequencies for each population were estimated and the squared product of the most common alleles obtained for each locus (Jones et al. 1998). The product across all loci was estimated to be the most likely genotype frequency in the population, and thus the highest probability of making a β-error during "fingerprint" analyses. The α-error here is essentially zero, but Jones et al. (1998) warn that *de novo* mutations can cause false exclusions. The resolution of individuals as unique was quantified for multiple locus combinations in each population, region, subspecies, and across *M. salmoides*. Locus combinations were used starting with the loci showing greatest resolution according to ANOVA variance components within individuals in populations.

RESULTS

Primers

Eleven of 45 original primer pairs amplified *M. salmoides* microsatellite loci and exhibited either suitable levels of polymorphism or distinctly differentiated allele frequencies between the subspecies (Table 3, Figure 2, Figure 3, Figure 4, Figure 5). The suite of 11 primers consisted of four primer pairs originally designed to amplify *Lepomis macrochirus*, four designed for *M. dolomieu* amplification, one for *L. auritus*, and two primers designed specifically for the amplification of *M. salmoides* microsatellite loci. Though most loci were amplified consistently in both subspecies, two of the loci, Lma010 and Ms025 consistently amplified alleles from *M. s. salmoides* samples but amplified little to no product in *M. s. floridanus*. Lma010 primer pairs also amplified what seemed to be a paralogous sequence, monomorphic in both subspecies, at 128 base pairs (bp), which was ignored during analyses, although this introduces a potential competitive priming problem given that excessive primer binding to this locus could prevent amplification at Lma010. Ms021 proved to be highly informative, but was the most difficult to score given a difference of a single motif repeat between the most common alleles, requiring the use of known samples to be run alongside the ladder in order to accurately score the products. The remaining loci did not pose any particular problems for amplification or scoring. Annealing temperatures ranged from 46.2 – 59.6°C using identical PCR cycles among all primer pairs except Ms021, which was used as described by DeWoody et al. (2000), with modifications (Table 3). The annealing

temperatures used deviated as much as 12.4° C from the literature, though in some there was no deviation.

Polymorphism, Hardy-Weinberg Equilibrium (HWE) and Linkage Disequilibrium

Given 14 populations and 11 loci there were 154 tests for HWE. The number of tests that returned significant results (23) was higher than expected due to type I error given the null hypothesis, that the expected and observed heterozygotes are equal, was true. Global tests for HWE across loci revealed significant heterozygote deficiencies in only the Lma010 and Ms025 loci. Heterozygote deficiencies in microsatellites suggested the presence of null alleles (Chakraborty *et al.* 1992), as did the lack of amplification in most individuals at these two loci in *M. s. floridanus*. Based on heterozygote deficits overall and over several allelic classes at these loci, the presence of null alleles was inferred and corrected genotypes were produced using *Micro-Checker 2.2.0* (http://www.microchecker.hull.ac.uk/). Following the correction for null alleles and sequential Bonferroni adjustments for multiple comparisons, no tests indicated significant departures from HWE (Table 2). Exact tests for linkage disequilibrium, at an α-level of 0.05, also showed no significant results after Bonferroni corrections.

Complete genotypes were scored at 11 loci for the 473 *M. salmoides* included in the analyses with no two individuals in any population exhibiting the same genotype. The 11 marker system utilized detected considerable variation across all loci and populations. Estimates of expected heterozygosity averaged from 0.175 for locus Mdo006 to 0.808 for locus Lma021, and the number of alleles per locus ranged from three at locus Mdo004 to 38 at locus Lma021 (Table 2; Table 3). Relative to previous

studies six of eleven loci exhibited lower levels of heterozygosity, two of eleven loci showed a reduced number of alleles per locus, and allelic size ranges were comparable with no dramatic deviations (Table 3). Within populations the Hillsboro River sample expressed the lowest degree of heterozygosity (mean ± σ: 0.307 ± 0.30), while Lake Fryer had the greatest (0.556 ± 0.25) (Table 2). Overall levels of expected heterozygosity were greater in *M. s. salmoides* populations (0.447 ± 0.28 vs. 0.364 ± 0.31), but *M. s. floridanus* populations expressed a greater number of alleles per locus (4.364 ± 4.35 vs. 3.886 ± 2.71). When partitioned into regional groups levels of heterozygosity and the number of alleles per locus were lowest in *M. s. salmoides* populations residing at northern latitudes (0.369 ± 0.26; 2.909 ± 1.51), followed by *M. s. floridanus* populations (0.364 ± 0.31; 4.364 ± 4.22), and *M. s. salmoides* populations residing at southern latitudes (0.507 ± 0.25; 4.568 ± 2.88).

Genetic Differentiation and Distance

Estimates of genetic variation among populations indicated significant structuring within the data set, whether expressed as estimates of F_{ST} under the infinite alleles model, or R_{ST} under the step-wise mutation model (Table 4). Differentiation between subspecies indicated a large degree of subdivision, producing a mean pairwise F_{ST} value of 0.573 ± 0.01 (mean ± SE), and a mean pairwise R_{ST} value of 0.871 ± 0.01. Additionally, genetic distances were analyzed in a Cavalli-Sforza and Edward's distance matrix which was visualized as an unrooted 1000 bootstrap UPGMA consensus tree (Figure 6) and with multidimensional scaling techniques (Figure 7). In either case distinct differences were

observed and there was 100% bootstrap support between subspecies using the UPGMA method, and a mean pairwise genetic distance of 0.394 ± 0.01 among these populations.

Differentiation was significantly less within subspecies groupings, particularly within *M. s. floridanus* (Table 4, Figure 6). The mean pairwise F_{ST} value was higher within *M. s. salmoides* (0.237 ± 0.02) compared to within *M. s. floridanus* (0.049 ± 0.01), and mean Cavalli-Sforza and Edwards distances between populations within subspecies concurred with this result (0.151 ± 0.01 vs. 0.031 ± 0.01 respectively). Among *M. s. salmoides* populations all but 6 of 36 pairwise F_{ST} and R_{ST} values were significant. This relatively large degree of genetic variation among *M. s. salmoides* populations supports the possibility of regional subdivision within the subspecies. Pairwise F_{ST} values indicated distinct groups of Pikes, Minnetonka, Pepin and Lake Charlotte from the Texas reservoir and river populations of Fryer, Kickapoo, Twin Oaks, and Marvin Lakes, and the Devil's River, with the hatchery and Florida populations remaining distinct. Pairwise R_{ST} values were similar, but indicated that the Lake Charlotte population was less differentiated from the group containing populations south of 37°N latitude. Private or virtually private (present at a ratio of > 10:1 between regional groups) alleles were identified at every locus.

Analyses of populations within regions indicated that the Devil's River population was distinct in its degree of genetic differentiation and distance from remaining *M. s. salmoides*. Measures of F_{ST} and genetic distance indicated that the Devil's River population was significantly different from the remaining *M. s. salmoides* populations and measures of R_{ST} concurred. Additionally, private or virtually private alleles were identified at all loci.

Analysis of Variance

Analysis of variance for all loci, with populations grouped by subspecies based on *a priori* knowledge of subspecies status, indicated that a great deal of variation among the samples was accounted for with between subspecies variation (Figure 8; $F_{CT} = 0.462$, $P < 0.001$), while variance among populations within each subspecies accounted for only 10% of the total. Analysis of molecular variance applying the same division concurred ($\phi_{CT} = 0.844$, $P < 0.001$). Variation was allocated more parsimoniously among regional groups, as indicated by measures of genetic distance (# of groups = 3), accounting for 48% of among group variance and forming more homogenous groups with only 5% of the variation among these populations ($F_{CT} = 0.480$, $P < 0.001$). In the third ANOVA the Devil's River population accounted for 16% of the total genetic variation within *M. s. salmoides*.

Additionally, locus-by-locus ANOVA's indicated that the loci Lma012, Mdo004, Mdo006, and Ms021 accounted for better than 50% of subspecies level variation per locus (Figure 9). Alleles present at ratios of greater than 10:1 were identified at all loci for subspecies comparisons (Table 5; Figures 2 through 5). Loci Lma012, Mdo004, Mdo006, and Ms021 also accounted for better than 50% of variation among regional groups (Figure 9). Most of the variation among individuals was accounted for at the loci Lma021 and Lar007 within *M. salmoides* (Figure 9).

Specifically looking at population level variation, 64% of the genetic variation at Mdo006 was accounted for between the Devil's River and remaining *M. s. salmoides* populations, followed by Mdo003 at nearly 37%, and Mdo007 at 28%. Alleles 135 bp at

locus Mdo003, 187 bp at Mdo007, and 160 bp at Mdo006 were present in the Devil's River population at substantially higher frequencies than in other populations, and allele 166 bp at Mdo006 was disproportionately low in relative frequency (Table 5).

Cluster Analysis and Assignment Tests

Structure analyses from k = 1 to 15 indicated that three clusters best accounted for variation in the data set according to the principles of HWE (Figure 10), though k = 2 (i.e. subspecies division) accounted for a large proportion of the variation in the data set. Bayesian analysis of *M. salmoides* populations at k = 2 produced clustering probabilities of 0.996 ± 0.01 (cluster 1) and 0.994 ± 0.02 (cluster 2) (Table 6). Beginning with the most discriminatory loci, and using from four to eleven loci, populations formed two distinct groups where no individuals were misclassified according to their assumed subspecies status and only 4% were rejected from their cluster at the highest stringency threshold. Results of *Arlequin's* assignment test concurred, indicating a strong division between subspecies using eleven loci, with mean log-likelihood ratio's of 6.3×10^{38} and 3.8×10^{38}. Based on *a priori* knowledge of subspecies status, no individuals produced greater likelihood values for clustering with individuals of the opposite subspecies (Table 6, Figure 11).

Individuals from the Lake Kissimmee and Lake Pepin samples were then added to the model one at a time, in a step-wise manner, in order to cross-validate the model. All individuals were correctly assigned to their assumed subspecies origin and cluster, with mean clustering probabilities of 0.996 ± 0.005 for *M. s. floridanus* and 0.996 ± 0.003 (0.004 ± 0.003 in cluster 1) for *M. s. salmoides*. No individuals were subsequently

rejected from their clusters at either the 0.95 or 0.99 threshold level (Table 6, Figure 12). The log-likelihood method corroborated these results with no individuals assigned to a group incorrectly (Table 6, Figure 11).

Although pure subspecies were easily resolved from each other using four to eleven loci in the sampled populations, the presence of introgression among wild populations and induced introgression resulting from stocking events, complicates the identification of subspecies. Using eleven locus genotypes, simulated intergrades were produced at several levels using the program *Hybridlab*. The intergrade analyses were also performed at k = 2 and assessed on a scale from 1 (*M. s. floridanus*) to 0 (*M. s. salmoides*). Bayesian analysis of intergrade populations showed that mean clustering probabilities approached those of pure subspecies by the third parental backcross generation (0.974 ± 0.046 and 0.025 ± 0.024) with nearly 80% of individuals classified within the range of the mean ± 3 standard deviations of either group after z-score transformations (Figure 13, Figure 14). It also indicated that F_1's (0.500 ± 0.052) and the first parental backcross generations (0.759 ± 0.106 and 0.230 ± 0.096) were readily distinguishable from each other and the other levels of introgression explored. Pure *M. s. floridanus* individuals produced clustering probabilities that were not shared with individuals within F_1 populations, though they did share clustering values with 6% of 1st generation backcrosses, 41% of 2nd generation backcrosses, and 67% of 3rd generation backcrosses to *M. s. floridanus* populations. Clustering probabilities were shared similarly within *M. s. salmoides* populations. Among adjacent intergrade populations the proportions of shared clustering values were highest between 2nd and 3rd generation backcrosses and lowest between F_1 and 1st generation backcrosses (Figure 14).

Bayesian inference at the regional level produced clusters as predicted from measures of differentiation and genetic distance. Populations originating in Florida, south of 37°N latitude (excluding the hatchery population), and north of 37°N latitude (including Lake Charlotte), formed distinct clusters. Mean clustering probabilities were highest for *M. s. floridanus* individuals (mean ± standard deviation: 0.992 ± 0.016), followed by individuals north (0.980 ± 0.038) and south of 37°N latitude (0.941 ± 0.155). Among all clusters only 1% of individuals were incorrectly assigned to a region of origin (Table 6,). Assessments of regional structure in *M. salmoides* using *Arlequin's* log-likelihood statistic offered contradictory results to the Bayesian analysis and concurred with R_{ST} results, indicating that 90% of Lake Charlotte individuals were more likely to have originated south of 37°N latitude.

Additional structuring within subspecies and regions was inferred with Bayesian analyses of k = 1 to 9 within *M. s. salmoides* populations, where the Devil's River sample formed a separate cluster. The Devil's River population was most pronounced in its division from other *M. s. salmoides* populations, with a mean clustering probability of 0.984 ± 0.012 at k = 3 and no individuals incorrectly classified, though 24% were rejected from the cluster at the highest stringency threshold (Table 6). Subsequently, Lake Charlotte fish and the hatchery fish formed separate clusters, though 7% of the hatchery fish were misclassified as wild *M. s. floridanus*, and most individuals did not show high clustering probabilities. Overall clustering solutions of $k \geq 5$ continued to form clusters without significantly increasing the models likelihood, indicating the remaining clusters were somewhat homogenous. Using Arlequin's classical assignment test again all individuals from the Devil's River sample were correctly classified.

Genetic Fingerprints

The squared product of the most common alleles across all loci indicated that β-errors could be expected at a mean rate of 1.4×10^{-4} within populations. The highest rate seen was 2.8×10^{-3} for the Hillsboro River population and the lowest was 2.7×10^{-7} for Lake Kickapoo. Given a free exchange of alleles within regions the rate of expected β-errors decreases to 7.3×10^{-6}, within subspecies it drops to 6.3×10^{-6}, and globally it is at 6.8×10^{-10}. Most of the variation among individuals is accounted for at the loci Lma021, Lar007, Mdo003, and Ms021 within *M. s. floridanus*, yet variation among individuals within *M. s. salmoides* was expressed (greatest allele frequency ≤ 0.699) at all loci with the exception of Ms021, Lar007, and Lma012 (Table 5). Given that few highly polymorphic loci were identified, the increase in resolution with loci is asymptotic (Figure 15).

DISCUSSION

Loci

Eleven microsatellite loci capable of resolving population structure and identifying subspecies, regional, and, in some cases, population origins of individual *M. salmoides* were identified from a pool of previously published microsatellite loci. Because loci were specifically selected for states of polymorphism and conserved differences among assemblages these findings are not directly comparable to other genetic studies, although estimates of heterozygosity did not deviate from levels seen in most other freshwater fishes (DeWoody and Avise 2000). Indeed, rather than the predicted decrease in polymorphism and allelic size with phylogenetic distance several of these loci recovered greater levels of heterozygosity and larger allelic sizes relative to the original studies. Though, differing sampling locations could explain this given the variance in levels of heterozygosity and genetic diversity among populations within this study.

The loci chosen had previously been used for studies in the family Centrarchidae assessing the evolution of microsatellites within the family, parentage analysis, understanding reproductive tactics, and quantifying mating success. Most were not highly polymorphic within the species in which they were originally identified. Levels of polymorphism were most pronounced at Lma021 and Lar007 in *M. salmoides*, which is consistent with the findings of Neff et al. (1999) and DeWoody et al. (2000), and moderate to low levels of polymorphism found at the remaining loci were consistent with

findings for these loci in other Centrarchid species (Colbourne et al. 1996; DeWoody et al. 2000; Neff et al. 1999). For the purposes of identifying hierarchical origins of individuals, conserved differences between and among groups of taxa, regions, or populations were particularly useful, but the presence of additional highly polymorphic loci would be ideal for identifying individuals uniquely (Angers et al. 1995; Jones et al. 1998; O'Reilly et al. 1996).

Population Structure

This is the first study to apply microsatellite loci toward resolving *M. salmoides* population structure, although previous attempts to resolve subspecies achieved some efficiency and reliability using other methods (Nedbal and Philipp 1994; Philipp et al. 1983; Williams et al. 1998; Williamson et al. 1986). The overall amount of genetic differentiation among all pairwise localities (mean pairwise F_{ST} = 0.383) indicated significant population structure in *M. salmoides*. Additional measures of differentiation (mean pairwise R_{ST}= 0.521) and distance (mean pairwise Cavalli-Sforza and Edwards distance = 0.258) concurred. Based on AMOVA and *Structure* analyses using multiple models the large degree of subdivision in the data set corresponds to existing subspecies divisions, accounting for a large proportion of the variation among localities. Indeed, *Structure* (Pritchard et al. 2000) analyses indicated that a division of populations into two inferred clusters accounted for 46% of the genetic variation among localities, and the composition of these clusters was identical to *a priori* assumptions of the subspecific status of individuals (Figure 10). Though, partitioning individuals to three geographic regions best accounted for genetic variation in the data set.

Within and among the localities sampled, *M. s. floridanus* expressed lower levels of genetic diversity than *M. s. salmoides*. These results contrast with those of Nedbal and Philipp (1994), who found greater mtDNA sequence divergence within *M. s. floridanus* than within *M. s. salmoides*. Philipp et al. (1983), on the other hand, found that intra-subspecies variation was lowest among *M. s. floridanus* samples for allozyme analyses. Smaller measures of genetic differentiation and distance among localities within *M. s. floridanus* may be due to the isolated geographic area from which they exist relative to the much larger range from which *M. s. salmoides* was sampled (Figure 1). However, even across geographic regions of comparable size *M. s. floridanus* showed lower mean levels of genetic differentiation and distance (mean pairwise $F_{ST} = 0.046$), quite possibly a consequence of a bottleneck event. Hypotheses include a smaller Pliocene refuge during what Near et al. (2003) described as fragmented tributaries or range contraction after Pleistocene glacial maximums. Near term Holocene climatic shifts and river basin changes may also have affected the structure subsequently revealed within each of the subspecies.

Pairwise measures of genetic differentiation and distance among localities within subspecies, also suggested significant structure within *M. s. salmoides* (mean pairwise F_{ST} among *M. s. salmoides* localities = 0.235). The presence of structure within *M. s. salmoides* was resolved as regional differences in allele frequencies for populations originating north and south of 37°N latitude, with some ambiguity as to which region individuals from Lake Charlotte were more genetically similar. These findings are in general agreement with Nedbal and Philipp (1994) who also found a distinction between their northernmost population and remaining *M. s. salmoides* populations, and identified

separate Pleistocene glacial refuges, as described by Bernatchez and Dodson (1991), and subsequent range expansions as the most probable cause of this variation. Thus, populations residing at northern latitudes may be the result of relatively recent lateral range expansion from populations isolated on the east coast of the United States or vertical range expansion from southern refuges. Both methods of range expansion from Pleistocene refuges would explain the relatively low levels of genetic variation in northern latitude populations.

Bayesian inference, measures of F_{ST} and genetic distance (Figure 6, Figure 7) suggest that Lake Charlotte samples, which do not exist in the range of the Pleistocene ice sheet, are more similar to *M. s. salmoides* populations in areas impacted by the glaciation events. Yet, pariwise R_{ST}'s and log-likelihood statistics indicate that the group is more similar to populations south of 37°N latitude. The discrepancy among methods may be the result of many *M. s. salmoides* samples in this study existing at the polar ends of 15° (~ 30°N to 45°N) of latitude, with Lake Charlotte residing at a somewhat intermediate location (37°N). Differential findings among methods may also be the result of regional intra-subspecific introgression in *M. s. salmoides*. Philipp et al. (1981, 1983) have previously identified a genetic gradient along this transect in *M. salmoides* populations for several allozymes, and genetic gradient may be present for these microsatellite loci with latitude. Microsatellites though, are not subject to the selective pressures that enzymatic loci experience, whose allele frequencies are commonly correlated with their thermal environment (Hines et al. 1983; Philipp et al. 1985; Yardley et al. 1974). Thus, these results might imply either geographically isolated groups or a

genetic gradient along this transect, but additional intermediate populations need to be sampled to resolve this.

Regional groups were designated *a priori* and then examined for additional structure. Based on measures of genetic differentiation and distance, samples from the Devil's River were clearly differentiated from the remaining *M. s. salmoides* populations and their inferred region of origin. Differences in allele frequencies in the Devil's River are particularly interesting, given that this was the only lotic population sampled within *M. s. salmoides*. In comparison, the Hillsboro River population (the only lotic *M. s. floridanus* population) was similar in allele frequency to other *M. s. floridanus* populations, which were more similar overall according to measures of genetic distance and differentiation among populations within subspecies. The Devil's River population also may be morphologically distinct among *M. salmoides* according to Dr. C. Hubbs (personal communication) and Edwards (1980), indicating that it may be endemic to this drainage. Though, the presence of introduced *M. dolomieu* populations and their ability to hybridize with *M. s. salmoides* (Beaty and Childers 1980; Whitmore and Hellier 1988) suggest that differences in allele frequencies may be due to this alone. The fact that most of the genetic variation between the Devil's River and remaining populations exists at loci previously isolated in *M. dolomieu* supports this conclusion. However, the Devil's River exist in a drainage basin with a prevalent endemic fish fauna, thus with out additional clarification the existence of endemism or introgression is not resolved.

Assignment Success

Assignment tests, such as the ones implemented in *Structure* and *Arlequin*, are powerful tools for determining the origins of unknown individuals, and are increasingly employed for genetic assessment, forensics, and the practical management of fish populations (Perez-Enriquez et al. 1999; Primmer et al. 2000; Roques et al. 1999; Shaw et al. 1999; Was and Wenne 2002). The correct assignment of 100% of sample individuals to the appropriate subspecies of origin, using either of these assignment methods, demonstrates that microsatellite markers can be used to resolve identities of *M. salmoides* subspecies status both at the population and individual level.

Furthermore, analysis of simulated intergrades also supports the claim that the extent of subspecific allelic contributions to populations and individuals can be estimated, but that multiple parental backcrosses quickly eliminate the ability to resolve subspecies status. For introgressed populations individual z-scores at each suspected level of introgression (previously simulated populations) would provide an exact probability of belonging to each distribution of probability values based on the area under the curve. It would be prudent to apply this method to suspected intergrades.

Estimation of subspecific status using microsatellites offers more resolution than the published allozyme analyses. Although no locus is fixed differentially between subspecies, private and virtually private alleles are present between the subspecies, and eleven microsatellite markers have been quantified. For any individual sampled, particularly from an intergrade zone, the probability of incorrectly classifying fish to subspecies is greatly reduced because it can be assessed probabilistically across all loci and excluded based on the presence of one of many specific alleles. The level of

discrimination that is possible with microsatellite loci is substantially greater than with previous analyses. Allozyme analysis of two loci with fixed subspecific differences, for example, will incorrectly identify first generation parental backcrosses as consisting of only *M. s. floridanus* or *M. s. salmoides* alleles at a rate of 0.17 (J. Dixson, personal correspondence, see Figure 16) and identify them as F_1's at the same rate. Philipp et al. (1983) have previously cautioned against making inferences of an individual's subspecific composition based on two allozyme markers.

The assignment of individuals to regions of origin was nearly as accurate as subspecies inferences. Depending on the method, and assuming that Lake Charlotte individuals were assigned to the correct region by the method in use, 97% to 100% of individuals were correctly assigned to their region of origin. And, based on the clusters observed, and the concurring lines of evidence for assignment success, it is evident that allele frequencies in the Devil's River are distinct enough to be used as identifiable population level markers, particularly the loci Mdo007 and Mdo006 which exhibit alleles 187 bp and 160 bp in length at a higher frequency relative to the general population.

Genetic Fingerprints

Inadequate levels of polymorphism were found to enable uniquely fingerprinting individuals. While loci Lma021, Lar007, and Mdo003 offered some discriminatory power, using all eleven loci allowed for an estimated rate of β-error up to 2.8×10^{-3} within populations. Thus, theoretically, on average one in every 2800 individuals would show an identical genotype. This offers little discriminatory power and an inefficient number of loci compared to other studies of telosts which employed microsatellites (DeWoody et

al. 2000; Perez-Enriquez et al. 1999). This does not mean that microsatellites cannot be used for resolving individuals, but rather that this eleven marker suite is not powerful enough for this use. The additional identification of loci similar in discriminatory power to Lma021 and Lar007 is required before a microsatellite marker suite is used for individual fingerprints in *M. salmoides* populations.

Applications and Implications

Potter and Barton (1986) summarized the goals of stocking as the maintenance of quality fisheries, a diversity of fish species, and a healthy aquatic environment, while meeting the demand for angling. More specifically though, stocking and translocation events take place in order to create fisheries, to supplement systems under high angling pressure, and to aid in the recovery of fish stocks, and most importantly to improve public relations with citizens who utilize the resource. This can be seen in *M. salmoides* which is commonly used for this practice, and in fact, is one of the most commonly introduced game species in the United States (U.S. Congress 1993).

The consequences of introducing non-native species are well documented (Taylor et al. 1984). These include the direct displacement of native species, elimination of genetic integrity in isolated populations, homogenization of genomes which can ease barriers to disease, and the introduction of maladaptive alleles which may lead to an overall decrease in the fitness of individuals (i.e. outbreeding depression) (Philipp 1992). Additionally, hatchery stocks may commonly exhibit low levels of heterozygosity and genetic diversity due to a reduced effective population size and the resultant effects of inbreeding and genetic drift (Hansen and Loeschcke 1994). Although, hatchery

operations have made attempts to eliminate many of these deleterious effects through the production of triploid and other polyploid individuals, essentially eliminating stocked individuals from the native gene pool (Fries et al. 2003). It is less clear that hatcheries have accounted for the genetic diversity and consequent compatibility of stocks to the wild congeners.

Previous stocking efforts did not have access to an overall genetic structure for *M. salmoides* for use in assisting their efforts. The results provided here will help guide current and future fisheries management efforts. For example, these results suggest negative effects of introducing fertile *M. s. floridanus* into most wild populations. The relatively low levels of genetic distance, differentiation, and diversity observed within and among native populations in this taxon would inevitably be reduced when coupled with aquaculture production. Furthermore, the recognition of genetically distinct northern and southern latitude populations within *M. s. salmoides* suggest that transfers among regions should be as meticulously scrutinized, as any transfer of stocks between subspecies would be. As the cessation of translocation events is unlikely it would be prudent to confine these events within the regions outlined here, keeping in mind that some populations, (ie. the Devil's River), could eventually be recognized as unique and should be conserved until the issue is resolved. Thus, solutions to current practices might include the creation of regionally specific hatchery stocks, the integration of population genetic principles into the overall stocking designs, and incorporating informed management goals within those designs which are consistent with ensuring the integrity of unique genetic lineages.

Conclusion

In summary, the analyses indicate that at these eleven loci the subspecies are easily resolved from one another as a consequence of distinct differences in allele frequencies. It is recommended that microsatellites be used in place of allozyme markers as the primary tool for discriminating an individual's population, regional, or subspecies status in *M. salmoides*. Though the presence of intergrades complicates the certainty with which an individual's or population's subspecific status can be identified, this is not a problem unique to microsatellite analysis. A greater number of markers or complete allelic fixation at a comparable number of markers would improve resolution. Thus, it may also be practicable to use microsatellites in concert with fixed differences at allozyme or pseudogene loci to ensure confidence.

The presence of regional subdivision within *M. s. salmoides* is also resolved, and assignment success quantified, though it is cautioned that more data is required to clarify the nature of allelic changes by latitude. The resolution and identification of individuals along regional divisions within *M. s. salmoides* may be as imperative as identifying subspecies status, with far-reaching implications to management and conservation. The identification of these genetic groups should be accounted for in management strategies for the species, particularly in the case of translocation events that must be closely examined to ensure informed management decisions are meeting overall goals. The possibility of a genetic bottleneck for *M. s. floridanus* also suggests that the transfer of fish from this region may be especially detrimental to overall levels of genetic diversity elsewhere within the species. Within regions it appears that specific populations may be identified by their allele frequencies, as is the case with fish from the Devil's River. This

may represent introgression with introduced *M. dolomieu*, or the presence of a population endemic to this drainage within Texas.

APPENDIX A:
TABLES

Table 1. Sample locations, sample size, and *a priori* knowledge of subspecies status based on geographic location, stocking records, and previous publications (Kassler et al. 2002; Philipp et al. 1983).

Population sampled	Abbreviation	Subspecific status	Location (wgs84)	Date collected	Sample size
Hillsboro River, FL	HILL	*M s. floridanus*	26 28, -80.08	2000-2002	35
East Lake Tohopekaliga, FL	ELT	*M s floridanus*	28 17, -81 17	2000-2002	27
Lake Kissimmee, FL	KISS	*M s. floridanus*	27 96, -81 38	2000-2002	51
Lake Medard, FL	MED	*M s floridanus*	27.94, -82 29	2000-2002	34
Texas Hatchery Population (A E. Wood)	HCH	*M. s. floridanus*	San Marcos, Texas	2000-2002	72
Lake Kickapoo, TX	KICK	*M s salmoides*	33 66, -98 78	2000-2002	27
Lake Charlotte, TX	CHAR	*M s. salmoides*	36.75, -96 65	2000-2002	50
Twin Oaks Reservoir, TX	TWIN	*M s salmoides*	31.12, -96 28	2000-2002	31
Lake Fryer, TX	FRY	*M. s salmoides*	36 22, -100 64	2000-2002	30
Lake Marvin, TX	MAR	*M s. salmoides*	35 88, -100 19	2000-2002	5
Devils River, TX	DEV	*M s. salmoides*	31 62, -98 87	2000-2002	37
Lake Minnetonka, TX	MINN	*M s salmoides*	44 91, -95 98	2000-2002	27
Lake Pepin, TX	PEP	*M s salmoides*	44 43, -92 18	2000-2002	24
Pike Lake, TX	PIKE	*M s salmoides*	43 31, -100 64	2000-2002	23

Table 2. Genetic diversity indices for eleven microsatellite loci in the fourteen *M. salmoides* populations used in this study. For individual loci no departures from Hardy-Weinberg Equilibrium or linkage disequilibrium were observed in any population after applying sequential Bonferroni correction for multiple tests ($\alpha = 0.05$, 1023 permutations). Presented are the expected heterozygosity (H_E), observed heterozygosity (H_O), *P*-value for exact tests of HWE (P), and the observed number of alleles per locus per population (A). Population abbreviations are in parentheses.

		M. s. floridanus			*M. s. salmoides*									
Locus		Hillsboro River (HILL) (n = 35)	Lake Kissimmee (KISS) (n = 27)	East Lake Tohopekaliga (ELT) (n = 51)	Lake Medard (MED) (n = 34)	Texas Hatchery (HCH) (n = 72)	Lake Kickapoo (KICK) (n = 27)	Lake Charlotte (CHAR) (n = 50)	Twin Oaks Reservoir (TWIN) (n = 31)	Lake Fryer (FRY) (n = 30)	Lake Marvin (MAR) (n = 5)	Devils River (DEV) (n = 37)	Lake Minnetoka (MINN) (n = 27)	Lake Pepin (PEPN) (n = 24)	Pike Lake (PIKE) (n = 23)
Lma010*	H_E	0 000	0 070	0 000	0 000	0 000	0 730	0 377	0 708	0 679	0 000	0 629	0 469	0 586	0 718
	H_O	0 000	0 037	0 000	0 000	0 000	0 704	0 360	0 839	0 800	0 000	0 757	0 370	0 625	0 609
	P	1 000	1 000	1 000	1 000	1 000	0 798	0 695	0 645	0 699	1 000	0 585	0 199	1 000	0 032
	A	1 000	2 000	1 000	1 000	1 000	6 000	5 000	5 000	5 000	1 000	5 000	3 000	4 000	4 000
Lma012	H_E	0 300	0 505	0 530	0 592	0 391	0 142	0 097	0 240	0 000	0 066	0 240	0 465	0 547	0 167
	H_O	0 229	0 556	0 549	0 471	0 403	0 111	0 080	0 226	0 000	0 033	0 243	0 556	0 458	0 130
	P	0 181	0 156	0 437	0 310	0 762	1 000	1 000	1 000	1 000	1 000	1 000	0 209	0 232	1 000
	A	3 000	5 000	4 000	4 000	4 000	2 000	2 000	4 000	1 000	2 000	2 000	2 000	3 000	3 000
Lma021	H_E	0 810	0 900	0 916	0 917	0 872	0 875	0 856	0 854	0 863	0 778	0 838	0 589	0 568	0 645
	H_O	0 800	0 963	0 922	0 912	0 903	0 926	0 880	0 903	0 800	0 800	0 892	0 444	0 542	0 652
	P	0 135	0 988	0 468	0 059	0 120	0 597	0 416	0 826	0 324	0 874	0 819	0 248	0 561	1 000
	A	12 000	13 000	22 000	16 000	10 000	13 000	9 000	9 000	10 000	5 000	11 000	8 000	5 000	4 000
Lma120	H_E	0 288	0 297	0 000	0 193	0 217	0 750	0 722	0 362	0 545	0 533	0 591	0 453	0 661	0 580
	H_O	0 343	0 296	0 000	0 206	0 222	0 667	0 660	0 355	0 567	0 400	0 622	0 519	0 750	0 609
	P	0 557	1 000	1 000	1 000	1 000	0 156	0 486	1 000	0 160	1 000	0 654	0 667	0 379	0 294
	A	2 000	3 000	1 000	3 000	3 000	5 000	5 000	2 000	4 000	2 000	3 000	2 000	6 000	3 000
Lar007	H_E	0 651	0 908	0 803	0 870	0 849	0 440	0 133	0 603	0 641	0 733	0 654	0 000	0 000	0 000
	H_O	0 543	0 963	0 765	0 765	0 833	0 407	0 120	0 581	0 633	0 600	0 730	0 000	0 000	0 000
	P	0 061	0 853	0 365	0 480	0 274	0 730	1 000	0 514	0 680	1 000	0 100	1 000	1 000	1 000
	A	9 000	14 000	9 000	9 000	11 000	4 000	2 000	6 000	5 000	4 000	4 000	1 000	1 000	1 000
Mdo003	H_E	0 638	0 542	0 601	0 639	0 616	0 584	0 219	0 526	0 637	0 800	0 000	0 108	0 272	0 457
	H_O	0 657	0 556	0 667	0 471	0 611	0 630	0 180	0 548	0 733	0 800	0 000	0 074	0 167	0 435
	P	0 790	0 798	0 823	0 191	0 308	0 585	0 465	0 536	0 887	1 000	1 000	1 000	0 078	0 809
	A	4 000	3 000	3 000	3 000	5 000	5 000	3 000	4 000	4 000	3 000	1 000	2 000	5 000	3 000
Mdo004	H_E	0 000	0 000	0 147	0 000	0 028	0 528	0 632	0 063	0 412	0 711	0 431	0 528	0 534	0 571
	H_O	0 000	0 000	0 137	0 000	0 014	0 519	0 620	0 065	0 500	0 400	0 459	0 593	0 417	0 565
	P	1 000	1 000	1 000	1 000	1 000	0 887	0 671	1 000	0 603	1 000	0 692	0 457	0 622	1 000
	A	1 000	1 000	2 000	1 000	2 000	3 000	3 000	2 000	3 000	2 000	2 000	2 000	3 000	3 000
Mdo006	H_E	0 057	0 143	0 113	0 086	0 028	0 498	0 373	0 373	0 493	0 533	0 105	0 000	0 000	0 000
	H_O	0 029	0 111	0 059	0 059	0 014	0 630	0 300	0 419	0 533	0 400	0 081	0 000	0 000	0 000
	P	1 000	1 000	0 098	1 000	1 000	0 234	0 253	0 646	0 693	1 000	1 000	1 000	1 000	1 000
	A	2 000	3 000	2 000	2 000	2 000	2 000	2 000	2 000	2 000	2 000	2 000	1 000	1 000	1 000
Mdo007	H_E	0 161	0 262	0 000	0 300	0 583	0 833	0 625	0 849	0 762	0 778	0 385	0 372	0 507	0 559
	H_O	0 143	0 185	0 000	0 324	0 542	0 778	0 680	0 935	0 700	1 000	0 351	0 259	0 500	0 565
	P	1 000	0 358	1 000	0 563	0 517	0 530	0 503	0 708	0 704	1 000	1 000	0 135	1 000	0 431
	A	2 000	2 000	1 000	2 000	4 000	9 000	4 000	8 000	8 000	3 000	3 000	2 000	2 000	3 000
Ms021	H_E	0 467	0 592	0 477	0 532	0 560	0 000	0 000	0 000	0 288	0 000	0 000	0 000	0 000	0 000
	H_O	0 571	0 778	0 412	0 588	0 556	0 000	0 000	0 000	0 233	0 000	0 000	0 000	0 000	0 000
	P	0 311	0 183	0 379	0 746	0 578	1 000	1 000	1 000	0 501	1 000	1 000	1 000	1 000	1 000
	A	3 000	2 000	2 000	3 000	3 000	1 000	1 000	1 000	2 000	1 000	1 000	1 000	1 000	1 000
Ms025*	H_E	0 000	0 178	0 096	0 167	0 134	0 735	0 803	0 876	0 798	0 800	0 757	0 730	0 399	0 699
	H_O	0 000	0 148	0 078	0 176	0 125	0 630	0 780	0 742	0 700	0 800	0 730	0 778	0 417	0 609
	P	1 000	1 000	1 000	1 000	1 000	0 403	0 296	0 234	0 588	1 000	0 409	0 549	0 743	0 154
	A	1 000	4 000	4 000	3 000	5 000	6 000	9 000	11 000	9 000	4 000	4 000	5 000	5 000	5 000

*correction for null alleles

Table 3. Oligonucleotide sequences of primer pairs optimized for the amplification of *M. salmoides* microsatellite loci (prior to adaptation for M13(-29) IR label), the author, original species designation, annealing temperature (following adaptation for M13(-29) IR label, reported annealing temperature, expected heterozygosity (H_E), reported H_E, number and distribution of alleles, and reported number and distribution of alleles.

Primer	Primer Sequence 5'-->3' F/R	Species designation	Author	T_{ann}°C	Reported T_{ann}°C	H_E	Reported H_E	(Alleles) Size	Reported (Alleles) Size
Lma010	GTCTGTAAGTGTGTTTGCTG GAAACCCGAAACTTGTCTAG	Lepomis macrochirus	Colbourne et al 1995	57 7	50	0 36	0 56	(7) 139-151	(4) 117-134
Lma012	CTGCTCAGCATGGAGGCAG TTCTTCCACAATATTCTCGCC	Lepomis macrochirus	Colbourne et al 1995	45 6	58	0 31	0 27	(6) 124-154	(8) 89-135
Lma021	CAGCTCAATAGTTCTGTCAGG ACTACTGCTGAAGATATTGTAG	Lepomis macrochirus	Colbourne et al 1995	47 5	58	0 81	0 60	(38) 146-244	(6) 154-182
Lma120	TGTCCACCCAAACTTAAGCC TAAGCCCATTCCCAATTCTCC	Lepomis macrochirus	Neff et al 1999	59 6	50	0 44	0 56	(10) 204-224	(5) 211-245
Lar007	GTGCTAATAAAGGCTACTGTC TGTTCCCTTAATTGTTTTGA	Lepomis auritus	DeWoody et al 2000	47	47	0 52	0 92	(15) 121-193	(22) 121-209
Mdo003	AGGTGCTTTGCGCTACAAGT CTGCATGGCTGTTATGTTGG	Micropterus dolomieu	Malloy et al 2000	46 2	55	0 47	0 57	(9) 119-143	(6) 135*
Mdo004	TCTGAACAACTGCATTTAGACTG CTAATCCCAGGGCAAGACTG	Micropterus dolomieu	Malloy et al 2000	48 6	55	0 33	0 66	(3) 151-155	(3) 142*
Mdo006	TGAAATGTACGCCAGAGCAG TGTGTGGGTGTTTATGTGGG	Micropterus dolomieu	Malloy et al 2000	55	55	0 20	0 54	(4) 160-176	(3) 150*
Mdo007	TCAAACGCACCTTCACTGAC GTCACTCCCATCATGCTCCT	Micropterus dolomieu	Malloy et al 2000	53	55	0 50	0 10	(10) 179-205	(2) 172*
Ms021	CACTGTAAATGGCACCTGTGG GTTGTCAAGTCGTAGTCCGC	Micropterus salmoides	DeWoody et al 2001	58	58	0 20	N/A	(5) 214-224	(3) N/A
Ms025	CAATATTGCCAAAGCATC CATTTGATACTGAATTTATTG	Micropterus salmoides	DeWoody et al 2002	47 5	54	0 51	N/A	(16) 198-230	(8) N/A

*size of cloned sequence

Table 4. Population pairwise genetic differentiation, based on conventional F_{ST} (lower diagonal) and R_{ST} (upper diagonal) estimates for eleven microsatellite loci in fourteen *M. salmoides* populations. Abbreviations are according to Figure 1 and Table 1.

	HILL	KISS	ELT	MED	HCH	KICK	CHAR	TWIN	FRY	MAR	DEV	MINN	PEP	PIKE
HILL	*	0 064	*0 033	0 108	0 087	0.940	0 978	0.942	0 947	0 988	0 912	0 801	0 923	0 861
KISS	0 067	*	*0 002	*-0 008	*0.005	0 859	0 928	0 868	0.872	0.884	0.834	0.676	0 824	0 755
ELT	0 061	0 032	*	*0.022	*0.023	0 907	0 948	0 911	0 914	0.937	0.884	0.773	0 885	0 834
MED	0 065	0 015	0 035	*	*-0 005	0 866	0 929	0 874	0 878	0 894	0 841	0 694	0 834	0 773
HCH	0 082	0 024	0 069	0 040	*	0 888	0 928	0 893	0.895	0.912	0.868	0 761	0 864	0 823
KICK	0 565	0 502	0 553	0 509	0 521	*	0 232	*-0.008	*-0 004	*0.129	*0 082	0 259	0 223	0 226
CHAR	0 614	0 567	0 604	0 571	0 569	0.183	*	0.176	0 159	*0.115	0.417	0 507	0 546	0 249
TWIN	0 593	0 536	0 581	0.542	0 550	0.069	0 245	*	*-0 012	*0 083	*0 138	0 307	0 283	0 215
FRY	0 558	0 501	0 550	0 508	0.523	0 058	0 210	0 036	*	*0 084	*0 141	0 306	0 281	0 212
MAR	0 648	0 570	0 623	0 572	0 576	0.103	0 251	0 089	0.051	*	*0 312	0.359	0 446	*0 113
DEV	0 613	0 560	0 597	0 563	0.559	0 193	0 402	0 261	0 215	0 255	*	0 146	*0 065	0 346
MINN	0 641	0 584	0 622	0 578	0 571	0 310	0 226	0 360	0 335	0 373	0 478	*	*0 057	0 309
PEP	0 627	0 567	0 609	0 564	0 558	0 269	0 188	0.338	0 311	0 392	0 461	0 128	*	0 371
PIKE	0 634	0 574	0 617	0 574	0 572	0 208	0 142	0 253	0 227	0 240	0 419	0 152	0 106	*

*not significant following sequential Bonferroni corrections ($\alpha = 0.05$, 1023 permutations)

Table 5. Population and subspecies allele frequencies are reported for all loci (0 bp = null allele).

		M s floridanus					*M s salmoides*									*M s f*	*M s s*
locus	allele size (bp)	Hill	Kiss	Elt	Med	Hch	Kick	Char	Twin	Marv	Fry	Dev	Minn	Pep	Pike	(N = 219)	(N = 254)
Lma010*	0	1 000	0 981	1 000	1 000	1 000	0 276	0 034	0 218	0 000	0 215	0 503	0 721	0 583	0 000	0 996	0 283
	139	0 000	0 000	0 000	0 000	0 000	0 207	0 741	0 067	0 000	0 247	0 000	0 118	0 264	0 433	0 000	0 231
	141	0 000	0 000	0 000	0 000	0 000	0 057	0 000	0 120	0 000	0 017	0 014	0 000	0 110	0 033	0 000	0 039
	143	0 000	0 019	0 000	0 000	0 000	0 000	0 034	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 004	0 004
	145	0 000	0 000	0 000	0 000	0 000	0 423	0 155	0 461	1 000	0 452	0 385	0 161	0 000	0 267	0 000	0 367
	147	0 000	0 000	0 000	0 000	0 000	0 019	0 034	0 084	0 000	0 069	0 000	0 000	0 000	0 267	0 000	0 053
	149	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 014	0 000	0 043	0 000	0 000	0 006
	151	0 000	0 000	0 000	0 000	0 000	0 019	0 000	0 050	0 000	0 000	0 085	0 000	0 000	0 000	0 000	0 017
Lma012	124	0 826	0 704	0 675	0 542	0 788	0 000	0 000	0 063	0 000	0 000	0 000	0 000	0 000	0 000	0 716	0 008
	130	0 000	0 000	0 000	0 000	0 047	0 906	0 956	0 875	0 900	1 000	0 896	0 682	0 667	0 917	0 017	0 877
	134	0 022	0 167	0 163	0 333	0 098	0 094	0 044	0 042	0 100	0 000	0 000	0 318	0 250	0 056	0 147	0 082
	142	0 152	0 037	0 075	0 104	0 000	0 000	0 000	0 021	0 000	0 000	0 104	0 000	0 083	0 028	0 058	0 033
	148	0 000	0 019	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 003	0 001
	154	0 000	0 074	0 088	0 021	0 067	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 059	0 001
Lma021	146	0 000	0 000	0 000	0 000	0 000	0 024	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 002
	150	0 000	0 000	0 000	0 000	0 000	0 000	0 017	0 000	0 000	0 000	0 000	0 000	0 053	0 053	0 000	0 018
	152	0 000	0 000	0 000	0 000	0 000	0 048	0 069	0 105	0 000	0 034	0 000	0 000	0 000	0 000	0 000	0 039
	158	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 026	0 000	0 034	0 000	0 000	0 000	0 000	0 000	0 006
	164	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 020	0 000	0 000	0 000	0 000	0 002
	166	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 020	0 000	0 000	0 000	0 000	0 002
	170	0 000	0 000	0 000	0 000	0 000	0 071	0 000	0 000	0 000	0 000	0 020	0 000	0 000	0 000	0 000	0 016
	172	0 000	0 000	0 000	0 021	0 000	0 071	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 002	0 008
	174	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 020	0 000	0 000	0 000	0 000	0 002
	176	0 000	0 000	0 000	0 042	0 000	0 048	0 000	0 079	0 000	0 000	0 000	0 000	0 000	0 000	0 009	0 014
	178	0 000	0 000	0 000	0 000	0 000	0 262	0 000	0 263	0 000	0 293	0 200	0 000	0 026	0 000	0 000	0 131
	180	0 000	0 000	0 000	0 000	0 000	0 071	0 000	0 053	0 400	0 086	0 300	0 000	0 000	0 000	0 000	0 074
	182	0 042	0 114	0 069	0 042	0 188	0 000	0 000	0 000	0 000	0 000	0 060	0 000	0 000	0 000	0 091	0 008
	184	0 000	0 000	0 000	0 083	0 000	0 071	0 000	0 000	0 000	0 000	0 120	0 000	0 000	0 000	0 014	0 024
	186	0 083	0 114	0 121	0 083	0 156	0 000	0 000	0 132	0 000	0 000	0 040	0 026	0 000	0 000	0 130	0 029
	188	0 042	0 000	0 086	0 000	0 094	0 071	0 000	0 053	0 200	0 121	0 140	0 000	0 000	0 000	0 068	0 047
	190	0 021	0 000	0 017	0 167	0 031	0 095	0 155	0 158	0 200	0 207	0 060	0 526	0 289	0 447	0 048	0 219
	192	0 000	0 023	0 017	0 146	0 000	0 024	0 000	0 000	0 100	0 086	0 000	0 000	0 026	0 026	0 018	0 020
	194	0 000	0 000	0 017	0 000	0 000	0 000	0 069	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 012
	196	0 000	0 023	0 017	0 000	0 016	0 048	0 190	0 053	0 100	0 086	0 000	0 026	0 000	0 000	0 014	0 053
	198	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 026	0 000	0 000	0 000	0 002
	200	0 021	0 000	0 034	0 042	0 000	0 000	0 000	0 053	0 000	0 017	0 000	0 026	0 000	0 000	0 016	0 010
	202	0 000	0 000	0 000	0 000	0 000	0 000	0 052	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 020
	204	0 000	0 114	0 103	0 021	0 141	0 000	0 207	0 000	0 000	0 034	0 000	0 026	0 000	0 000	0 077	0 049
	206	0 042	0 045	0 017	0 063	0 156	0 095	0 207	0 026	0 000	0 000	0 000	0 316	0 605	0 474	0 075	0 184
	208	0 083	0 227	0 155	0 000	0 063	0 000	0 034	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 096	0 004
	210	0 167	0 045	0 017	0 042	0 109	0 000	0 000	0 000	0 000	0 000	0 000	0 026	0 000	0 000	0 092	0 001
	212	0 042	0 023	0 034	0 021	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 018	0 000
	214	0 063	0 068	0 000	0 042	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 023	0 000
	216	0 000	0 091	0 069	0 104	0 047	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 059	0 000
	218	0 000	0 091	0 034	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 027	0 000
	220	0 375	0 000	0 052	0 042	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 079	0 000
	224	0 000	0 023	0 034	0 042	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 014	0 000
	226	0 000	0 000	0 017	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 002	0 000
	232	0 021	0 000	0 017	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 005	0 000
	238	0 000	0 000	0 017	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 007	0 000
	242	0 000	0 000	0 034	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 007	0 000
	244	0 000	0 000	0 017	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 007	0 000

Table 5. (continued)

Lma120	204	0 000	0 000	0 000	0 000	0 000	0 050	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 004
	208	0 000	0 000	0 000	0 000	0 000	0 050	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 044
	210	0 000	0 100	0 000	0 100	0 000	0 600	0 088	0 800	0 800	0 571	0 500	0 333	0 200	0 500	0 023	0 401
	212	0 833	0 833	1 000	0 833	0 889	0 150	0 088	0 000	0 100	0 143	0 433	0 000	0 100	0 100	0 899	0 122
	214	0 167	0 067	0 000	0 067	0 000	0 000	0 000	0 000	0 100	0 250	0 067	0 000	0 000	0 000	0 039	0 042
	216	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 200	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 026
	218	0 000	0 000	0 000	0 000	0 000	0 150	0 206	0 000	0 000	0 000	0 000	0 000	0 033	0 000	0 000	0 065
	220	0 000	0 000	0 000	0 000	0 056	0 000	0 382	0 000	0 000	0 036	0 000	0 667	0 533	0 400	0 018	0 246
	222	0 000	0 000	0 000	0 000	0 056	0 000	0 235	0 000	0 000	0 000	0 000	0 000	0 100	0 000	0 019	0 048
	224	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 033	0 000	0 000	0 002
Lar007	121	0 000	0 019	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 002	0 000
	129	0 053	0 096	0 023	0 063	0 031	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 046	0 000
	133	0 053	0 019	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 012	0 000
	141	0 579	0 135	0 273	0 229	0 125	0 000	0 000	0 093	0 200	0 069	0 000	0 000	0 000	0 000	0 263	0 021
	145	0 026	0 192	0 068	0 104	0 125	0 759	0 938	0 593	0 400	0 586	0 500	1 000	1 000	1 000	0 109	0 773
	149	0 026	0 192	0 023	0 125	0 078	0 074	0 000	0 111	0 000	0 000	0 000	0 000	0 000	0 000	0 057	0 020
	153	0 053	0 038	0 045	0 063	0 000	0 000	0 000	0 037	0 100	0 017	0 059	0 000	0 000	0 000	0 032	0 022
	157	0 132	0 115	0 250	0 104	0 359	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 213	0 002
	161	0 000	0 019	0 000	0 000	0 000	0 037	0 000	0 093	0 300	0 224	0 324	0 000	0 000	0 000	0 005	0 095
	165	0 000	0 096	0 091	0 188	0 094	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 098	0 000
	169	0 053	0 154	0 205	0 083	0 047	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 100	0 000
	173	0 000	0 000	0 000	0 000	0 000	0 130	0 063	0 074	0 000	0 103	0 118	0 000	0 000	0 000	0 000	0 065
	177	0 000	0 019	0 000	0 042	0 063	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 032	0 000
	189	0 026	0 019	0 023	0 000	0 063	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 023	0 000
	193	0 000	0 000	0 000	0 000	0 016	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 005	0 000
Mdo003	119	0 000	0 000	0 000	0 000	0 000	0 019	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 002
	125	0 208	0 231	0 464	0 250	0 125	0 093	0 000	0 111	0 000	0 200	0 000	0 000	0 050	0 184	0 261	0 066
	127	0 000	0 000	0 000	0 000	0 094	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 021	0 001
	129	0 563	0 615	0 411	0 542	0 516	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 523	0 001
	133	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 025	0 000	0 000	0 002
	135	0 000	0 000	0 000	0 000	0 016	0 574	0 159	0 630	0 600	0 517	1 000	0 000	0 050	0 105	0 004	0 388
	137	0 188	0 154	0 125	0 208	0 250	0 019	0 023	0 074	0 100	0 100	0 000	0 000	0 000	0 000	0 181	0 028
	141	0 042	0 000	0 000	0 000	0 000	0 296	0 818	0 185	0 300	0 183	0 000	0 950	0 850	0 711	0 009	0 505
	143	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 050	0 025	0 000	0 000	0 006
Mdo004	151	1 000	1 000	0 925	1 000	0 989	0 125	0 250	0 036	0 000	0 100	0 000	0 364	0 533	0 500	0 980	0 218
	153	0 000	0 000	0 000	0 000	0 000	0 208	0 289	0 000	0 400	0 150	0 761	0 636	0 433	0 467	0 001	0 345
	155	0 000	0 000	0 075	0 000	0 011	0 667	0 461	0 964	0 600	0 750	0 239	0 000	0 033	0 033	0 019	0 438
Mdo006	160	0 000	0 019	0 000	0 000	0 000	0 404	0 229	0 205	0 500	0 350	0 958	0 000	0 000	0 000	0 004	0 311
	166	0 000	0 000	0 000	0 000	0 021	0 596	0 771	0 795	0 500	0 650	0 042	1 000	1 000	1 000	0 004	0 686
	172	0 981	0 942	0 938	0 979	0 979	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 969	0 002
	176	0 019	0 038	0 063	0 021	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 023	0 001
Mdo007	179	0 000	0 000	0 000	0 000	0 000	0 167	0 500	0 235	0 000	0 267	0 000	0 206	0 531	0 559	0 001	0 304
	185	0 907	0 889	1 000	0 838	0 611	0 063	0 000	0 029	0 000	0 017	0 000	0 000	0 000	0 000	0 810	0 012
	187	0 000	0 000	0 000	0 000	0 078	0 104	0 000	0 088	0 300	0 017	0 740	0 000	0 000	0 000	0 029	0 139
	189	0 093	0 111	0 000	0 162	0 277	0 063	0 000	0 000	0 000	0 017	0 000	0 000	0 000	0 000	0 139	0 009
	191	0 000	0 000	0 000	0 000	0 000	0 333	0 019	0 147	0 400	0 250	0 240	0 000	0 000	0 000	0 000	0 124
	195	0 000	0 000	0 000	0 000	0 000	0 021	0 111	0 029	0 000	0 033	0 000	0 794	0 469	0 382	0 001	0 200
	197	0 000	0 000	0 000	0 000	0 034	0 125	0 370	0 265	0 300	0 367	0 020	0 000	0 000	0 000	0 018	0 158
	199	0 000	0 000	0 000	0 000	0 000	0 104	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 059	0 000	0 028
	201	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 118	0 000	0 033	0 000	0 000	0 000	0 000	0 000	0 016
	205	0 000	0 000	0 000	0 000	0 000	0 021	0 000	0 088	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 010

Table 5. (continued)

Ms021	214	0 000	0 000	0 000	0 000	0 000	0 056	0 000	0 000	0 000	0 000	0 183	0 000	0 000	0 000	0 001	0 022
	216	0 000	0 000	0 000	0 000	0 031	0 928	1 000	1 000	1 000	1 000	0 817	1 000	1 000	1 000	0 013	0 975
	218	0 708	0 436	0 621	0 542	0 505	0 017	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 567	0 001
	222	0 271	0 564	0 379	0 438	0 453	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 404	0 001
	224	0 021	0 000	0 000	0 021	0 010	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 016	0 001
Ms025*	0	1 000	0 925	0 960	0 910	0 937	0 000	0 027	0 000	0 000	0 000	0 000	0 144	0 000	0 208	0 946	0 042
	198	0 000	0 000	0 020	0 045	0 007	0 000	0 000	0 023	0 000	0 000	0 000	0 000	0 000	0 000	0 014	0 003
	200	0 000	0 000	0 010	0 000	0 021	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 006	0 000
	204	0 000	0 000	0 000	0 045	0 000	0 000	0 041	0 000	0 100	0 000	0 000	0 000	0 067	0 000	0 009	0 023
	206	0 000	0 000	0 000	0 000	0 014	0 095	0 095	0 159	0 000	0 017	0 182	0 038	0 000	0 000	0 003	0 065
	208	0 000	0 000	0 000	0 000	0 000	0 000	0 068	0 068	0 000	0 350	0 000	0 000	0 033	0 000	0 000	0 058
	210	0 000	0 000	0 000	0 000	0 000	0 357	0 189	0 250	0 000	0 233	0 318	0 097	0 767	0 489	0 000	0 300
	212	0 000	0 000	0 000	0 000	0 000	0 071	0 189	0 068	0 400	0 000	0 000	0 491	0 067	0 140	0 000	0 159
	214	0 000	0 000	0 010	0 000	0 000	0 071	0 311	0 068	0 000	0 050	0 000	0 230	0 067	0 022	0 002	0 091
	216	0 000	0 038	0 000	0 000	0 000	0 381	0 014	0 227	0 300	0 183	0 273	0 000	0 000	0 140	0 008	0 169
	218	0 000	0 000	0 000	0 000	0 021	0 000	0 000	0 023	0 000	0 033	0 000	0 000	0 000	0 000	0 004	0 006
	220	0 000	0 000	0 000	0 000	0 000	0 024	0 068	0 000	0 100	0 033	0 000	0 000	0 000	0 000	0 000	0 025
	222	0 000	0 019	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 004	0 000
	224	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 068	0 000	0 083	0 000	0 000	0 000	0 000	0 000	0 017
	226	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 023	0 000	0 017	0 227	0 000	0 000	0 000	0 000	0 030
	228	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 023	0 000	0 000	0 000	0 000	0 000	0 000	0 000	0 003
	230	0 000	0 019	0 000	0 000	0 000	0 000	0 000	0 000	0 100	0 000	0 000	0 000	0 000	0 000	0 004	0 011

*corrected for null alleles using *Micro-Checker* 2.2.0

Table 6. Clustering probabilities (mean ± σ) and likelihood ratios for Bayesian (Pritchard et al. 2000) and log-likelihood (Schnieder et al. 2000) methods of assigning individuals to (a) subspecies of origin, (b) geographic region of origin (region 2: south of 37°N latitude; region 3: north of 37°N latitude), and (c) population of origin. The proportion of individuals incorrectly assigned to a group using each method (P_I), and the proportion of individuals whose probability of belonging to their inferred group is less than 0.95 ($P_{R\,0.95}$) and less than 0.99 ($P_{R\,0.99}$), according to *Structure* analysis are reported. Samples from Lake Kissimmee and Lake Pepin were left out of the original analyses and used to cross-validate the subspecies model.

		Bayesian (*Structure*)				Frequency (*Arlequin*)	
		probability (mean±σ)	P_I	$P_{R\,0.95}$	$P_{R\,0.99}$	log-liklihood ratio	P_I
(a)	*M. s. salmoides*	0.994±0.019	0.00	0.02	0.04	6.3×10^{38}	0.00
	Lake Pepin	0.996±0.003	0.00	0.00	0.00	3.5×10^{35}	0.00
	M. s. floridanus	0.996±0.013	0.00	0.01	0.04	3.8×10^{38}	0.00
	Lake Kissimmee	0.996±0.005	0.00	0.00	0.00	6.9×10^{38}	0.00
(b)	*M. s. floridanus* Region 1	0.992±0.016	0.00	0.02	0.10	1.6×10^{39}	0.00
	M. s. salmoides Region 2	0.941±0.155	0.03	0.15	0.35	9.1×10^{13}	0.03
	M. s. salmoides Region 3	0.980±0.038	0.00	0.08	0.22	2.9×10^{6}	0.00
(c)	Devils River	0.984±0.012	0.00	0.03	0.24	3.9×10^{8}	0.00

APPENDIX B:
FIGURES

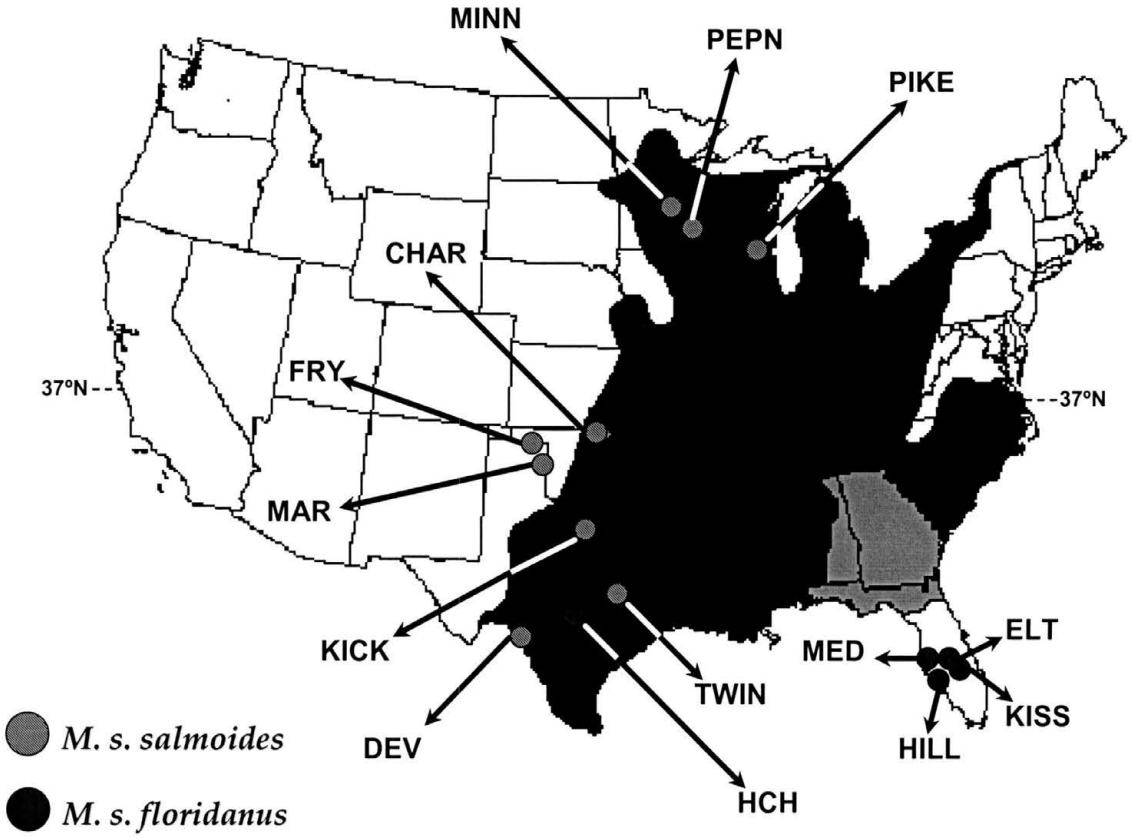

Figure 1. Map of the United States indicating the locations of *M. salmoides* populations sampled for this study. The populations are superimposed on their original ranges according to MacCrimmon and Robbins (1975) and Bailey and Hubbs (1949).

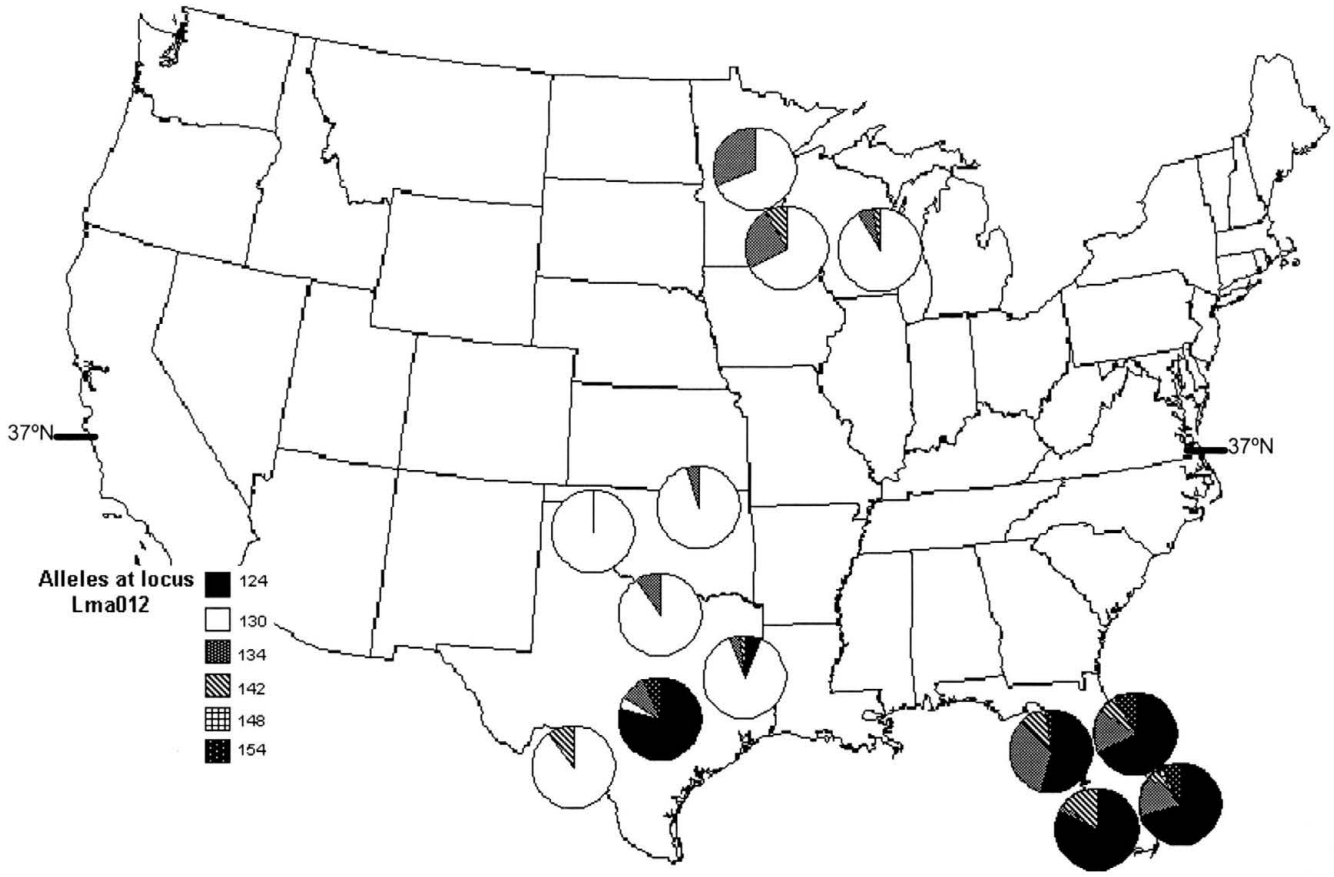

Figure 2. Map of the United States indicating the locations of *M. salmoides* populations included in this study. Presented are allele frequencies at the microsatellite locus Lma012. Pie charts indicate approximate allele frequencies for all 6 alleles.

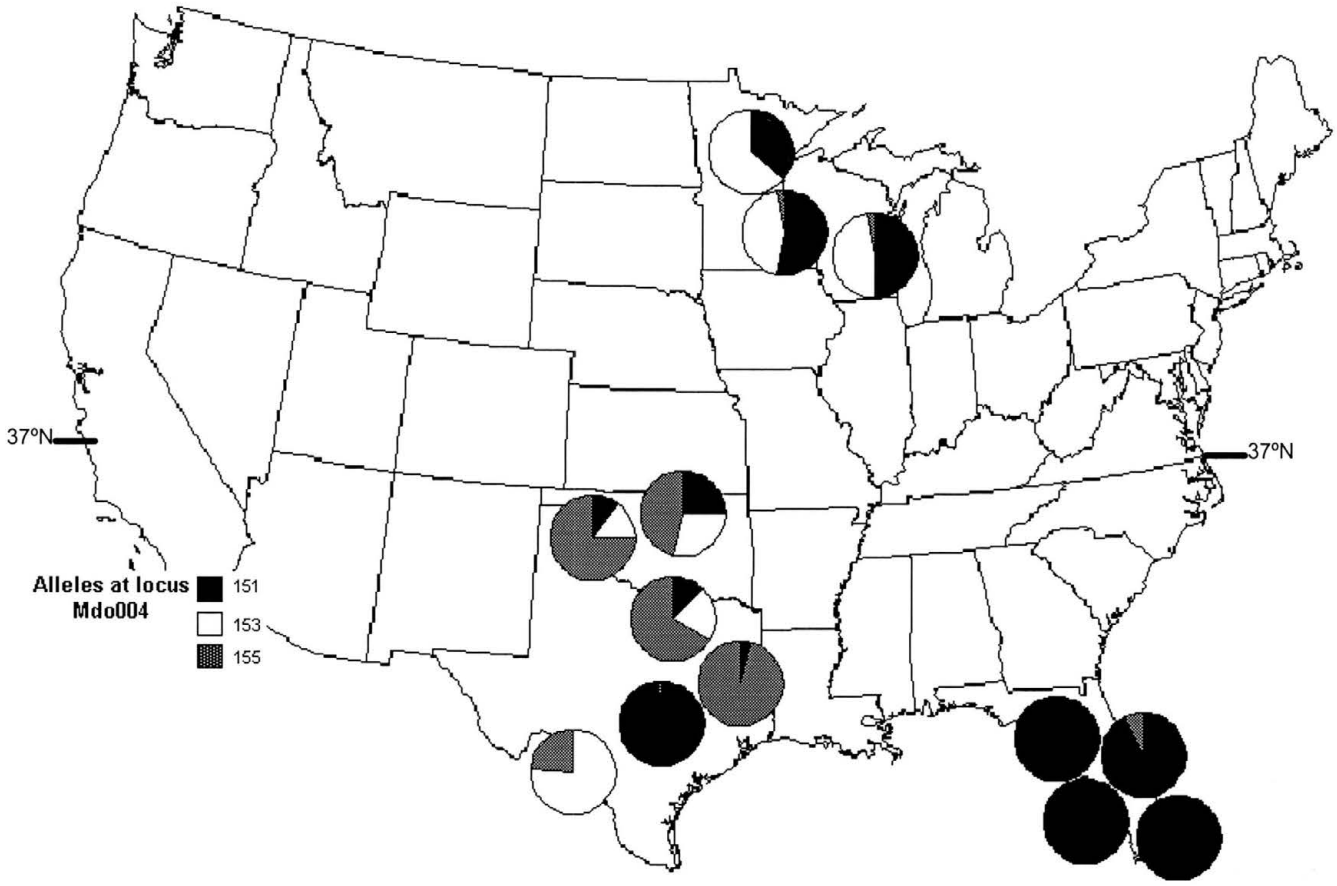

Figure 3. Map of the United States indicating the locations of *M. salmoides* populations included in this study. Presented are allele frequencies at the microsatellite locus Mdo004. Pie charts indicate approximate allele frequencies for all 3 alleles.

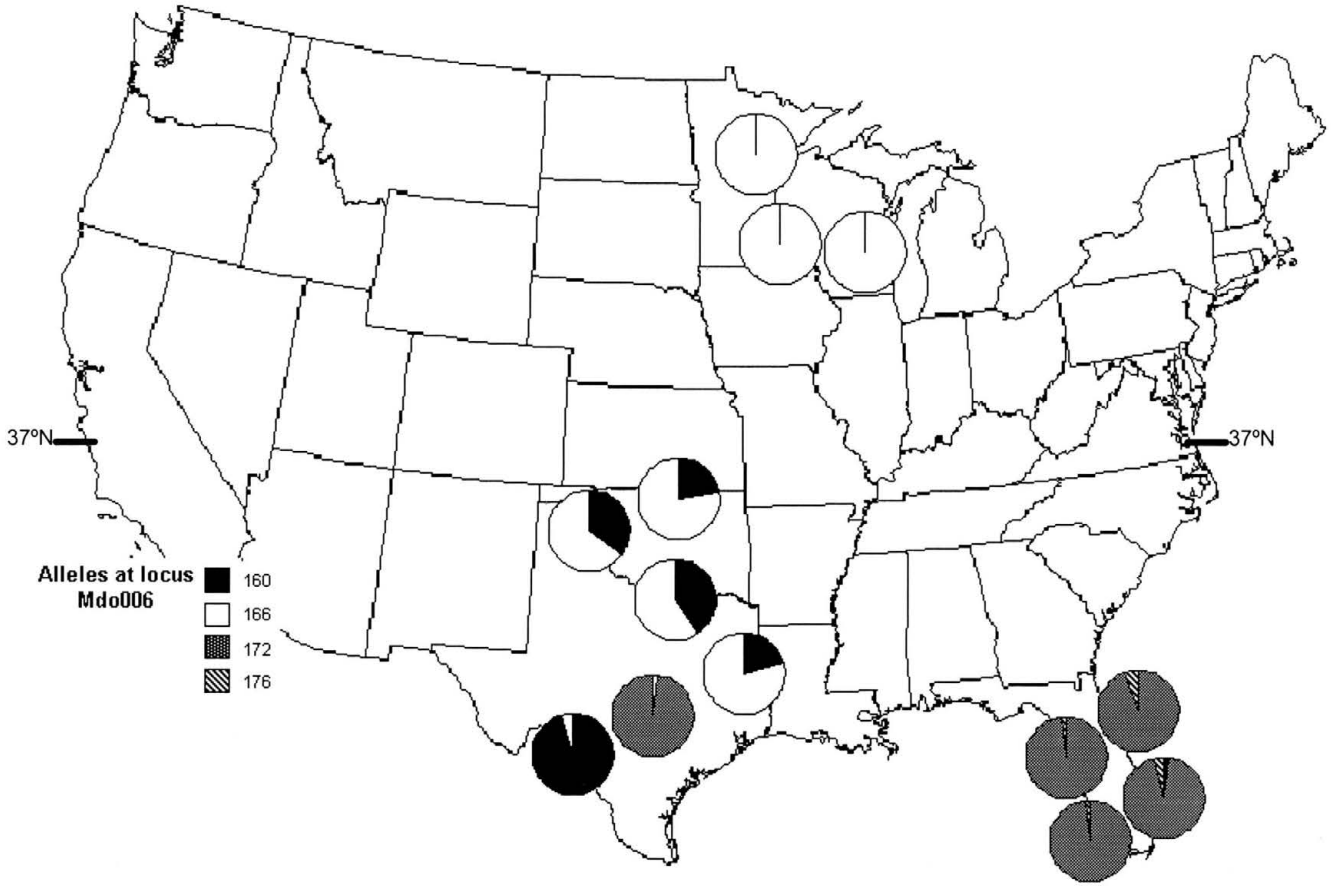

Figure 4. Map of the United States indicating the locations of *M. salmoides* populations included in this study. Presented are allele frequencies at the microsatellite locus Mdo006. Pie charts indicate approximate allele frequencies for all 4 alleles.

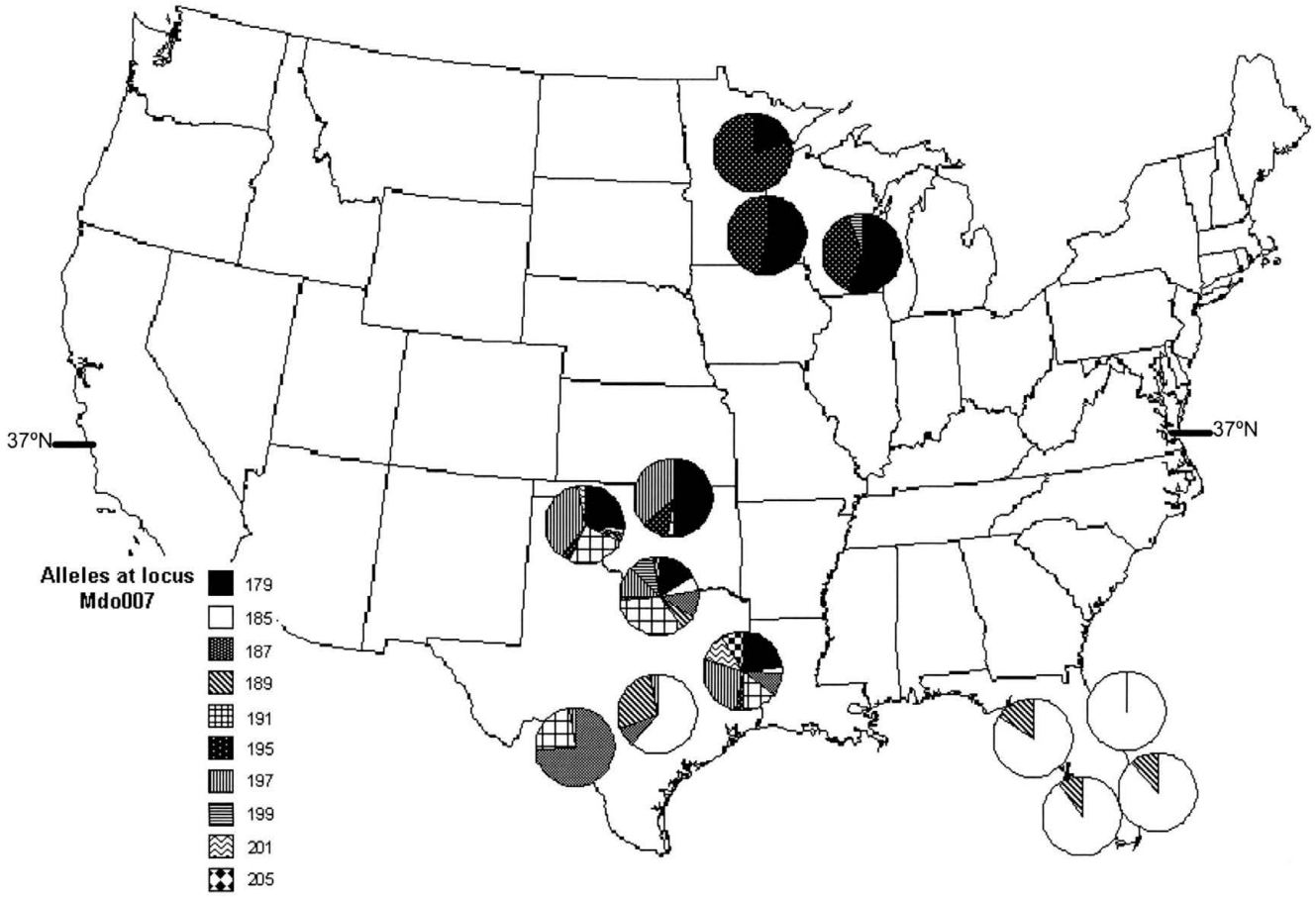

Figure 5. Map of the United States indicating the locations of *M. salmoides* populations included in this study. Presented are allele frequencies at the microsatellite locus Mdo007. Pie charts indicate approximate allele frequencies for all 10 alleles.

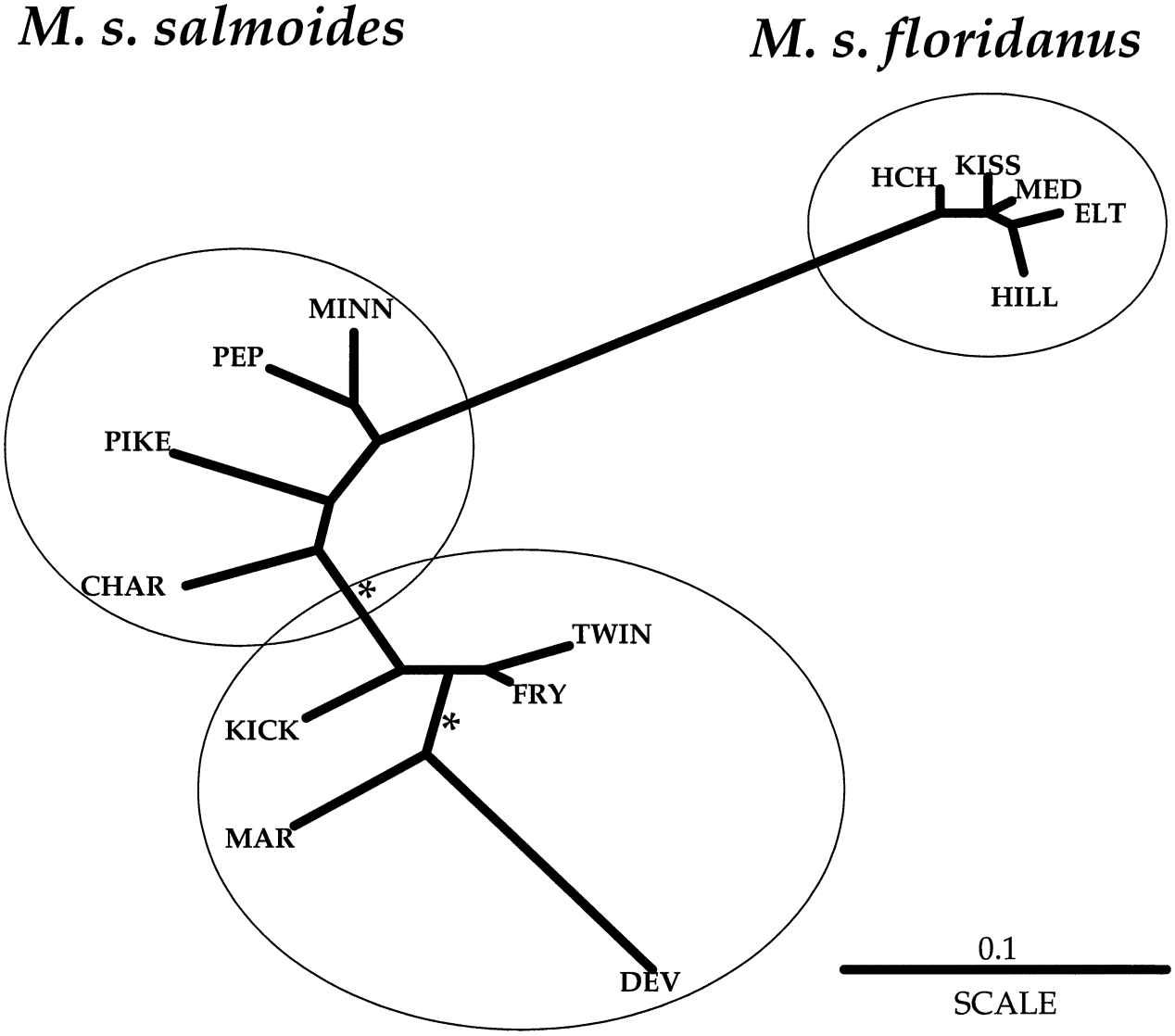

Figure 6. Unrooted UPGMA consensus tree for fourteen *M. salmoides* populations. Population distances are scaled based on Cavalli-Sforza and Edwards chord distance (Cavalli-Sforza and Edwards 1967) and * at nodes represent bootstrap support less than 50% (1000 bootstraps). Regional divisions are encircled below their subspecies status. Abbreviations are according to Figure 1 and Table 1.

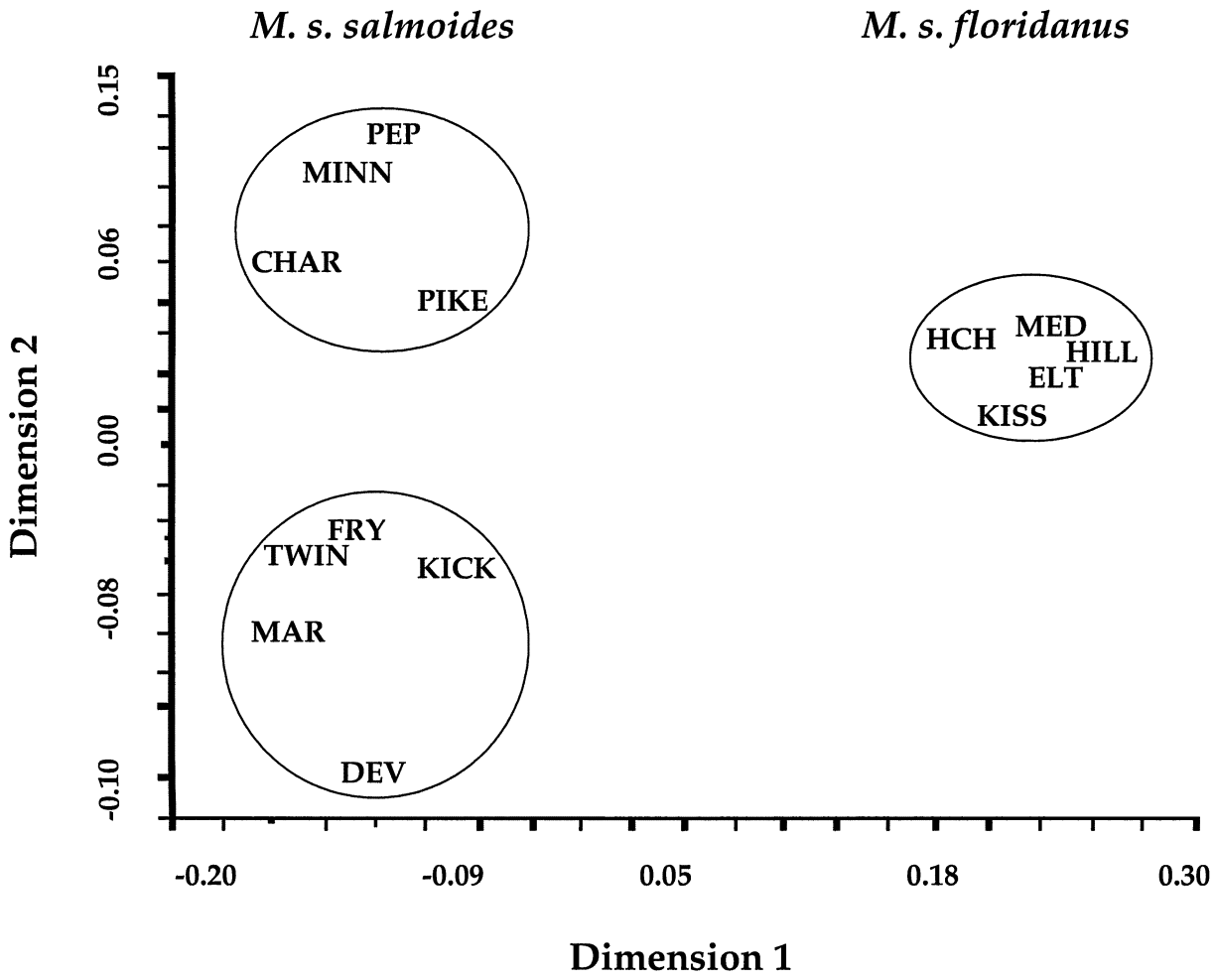

Figure 7. Multi-dimensional scaling (MDS) plots of nine *M. s. salmoides* populations and five *M. s. floridanus* populations based on Cavalli-Sforza and Edwards pairwise chord distances (Cavalli-Sforza and Edwards 1967). Regional subdivisions are encircled. Abbreviations are according to Figure 1 and Table 1.

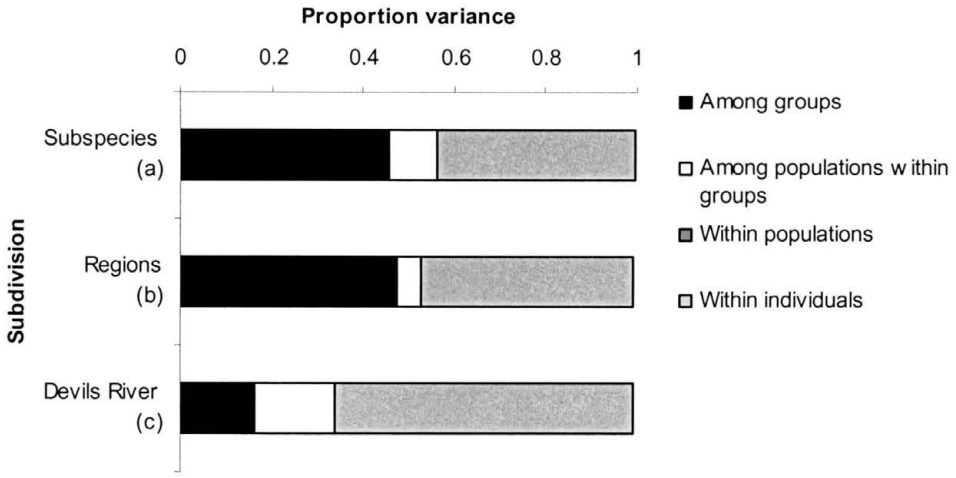

Figure 8. Hierarchical analysis of variance (ANOVA) based on genetic variation at eleven microsatellite loci in *M. salmoides*. (a) Results grouping populations by subspecies (*M. s. salmoides* vs. *M. s. floridanus*), (b) Results grouping populations by region, using three regions determined by population clustering analyses (Table 3, Figure 6, Figure 7): *M. s. floridanus* and *M. s. salmoides* populations north and south of 37°N latitude, (c) Results grouping *M. s. salmoides* into two groups identified from significant differences between the Devil's River and remaining *M. s. salmoides* populations.

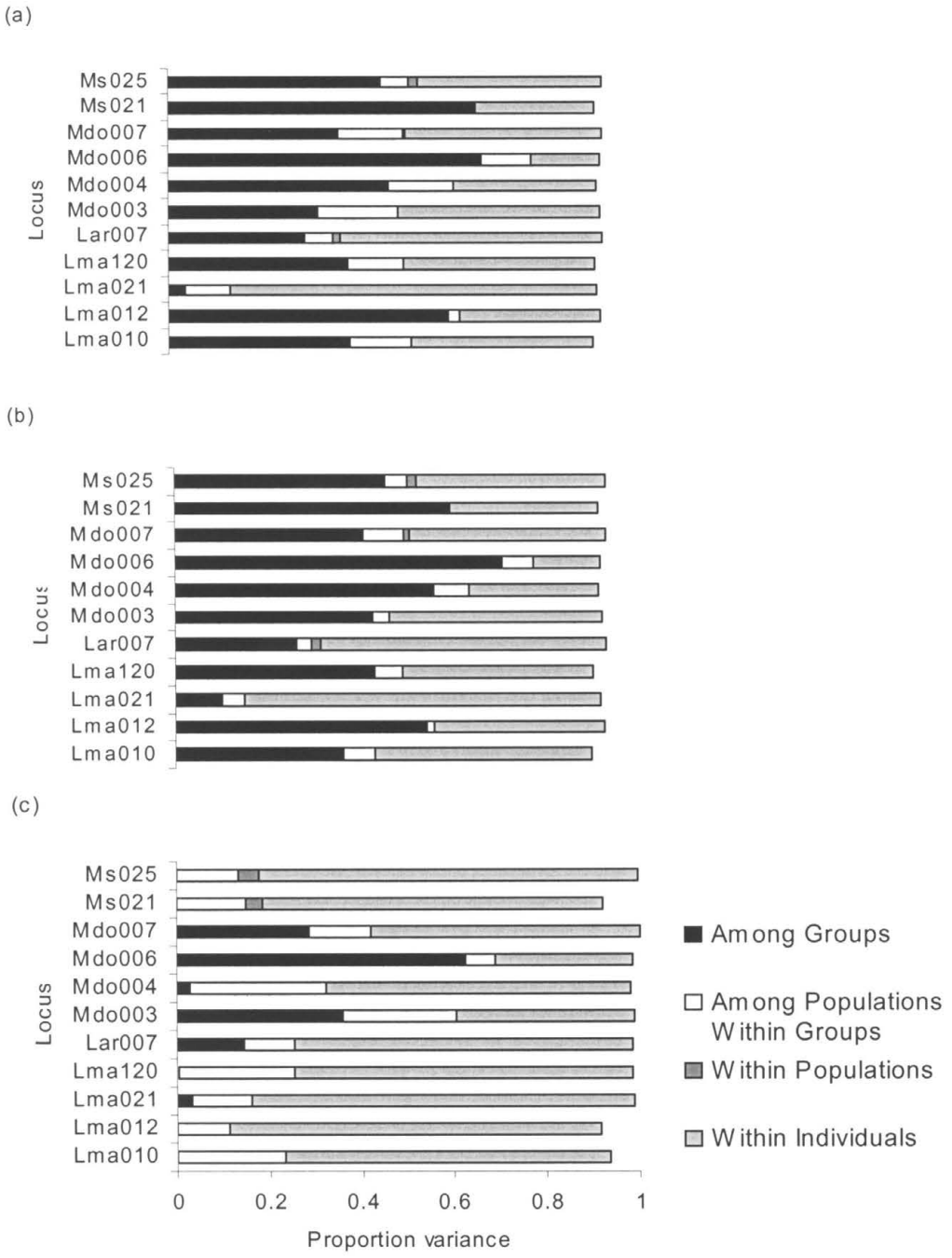

Figure 9. Locus-by-locus hierarchical analysis of variance performed to assess the differential contribution of specific loci to the resultant patterns of genetic variation at eleven microsatellite loci among and within groups and populations for the following divisions in *M. salmoides* (a) subspecies, (b) regions and (c) the Devil's River as described in Figure 8.

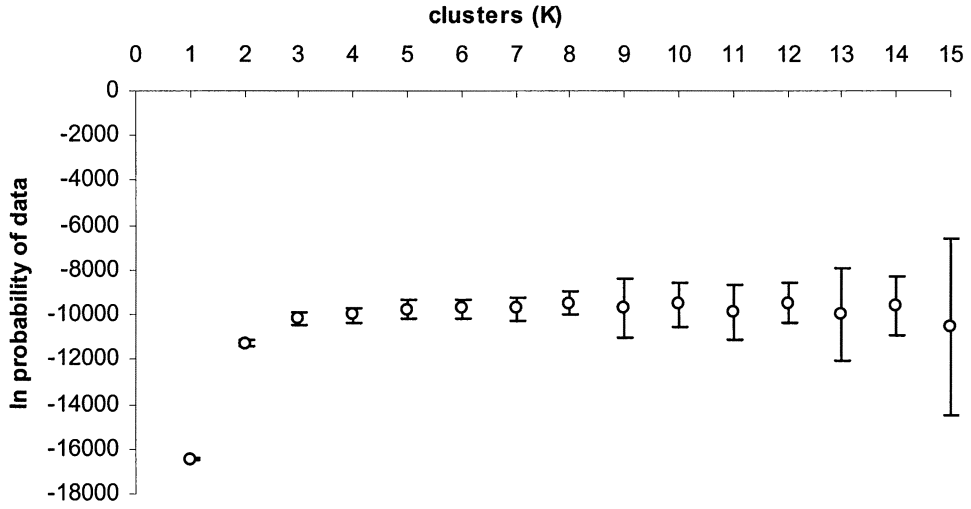

Figure 10. Clustering solutions (ln probability of data ± σ^2) describing the goodness of fit of the model (Bayesian inference of population structure) to the data set (multilocus genotypes for all individuals) with the purpose of identifying populations (clusters) in Hardy-Weinberg and linkage equilibrium. Clustering solutions of 2 – 3 account for most of the variation among populations.

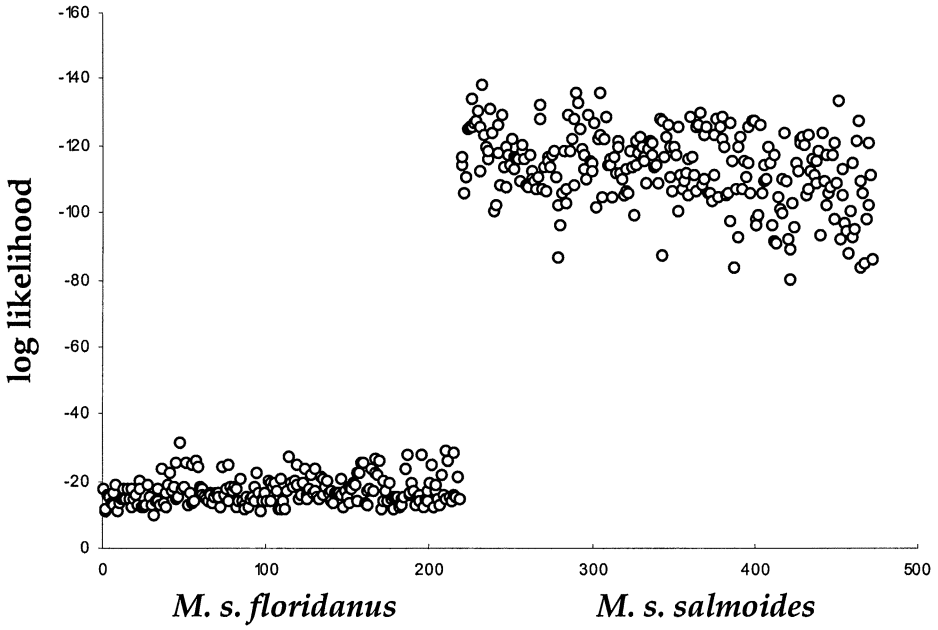

Figure 11. Log likelihood scores for origins in cluster 1, for *M. s. floridanus* ($N = 219$) and *M. s. salmoides* ($N = 254$) individuals, based on differences in allele frequencies for each subspecies at eleven microsatellite loci. All individuals were correctly assigned to their subspecies of origin.

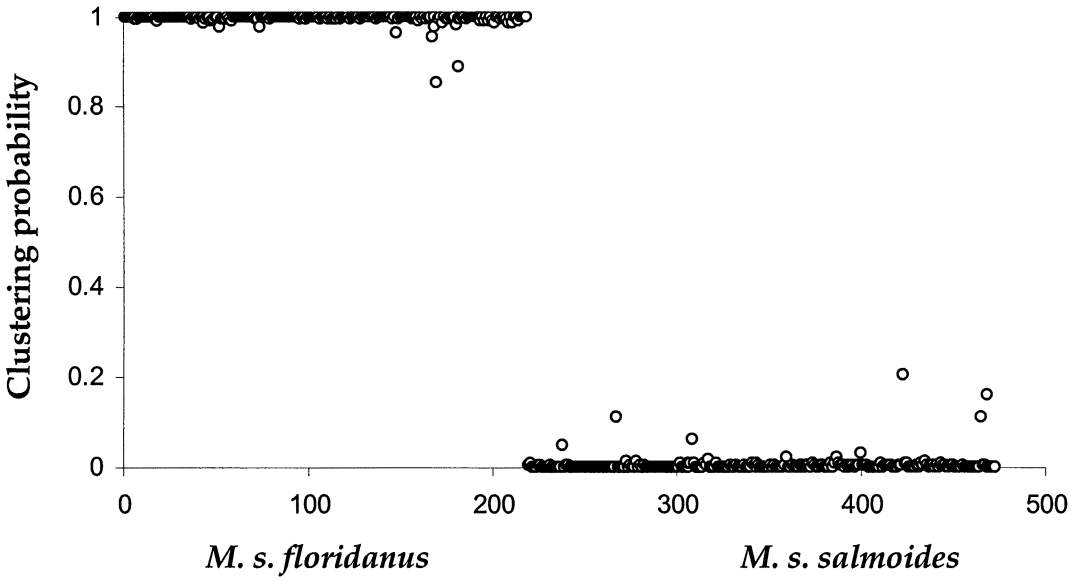

Figure 12. Bayesian determined clustering probabilities to cluster 1 based on genetic variation at eleven microsatellite loci in *M. s. floridanus* ($N = 219$) and *M. s. salmoides* ($N = 254$) individuals. All individuals were correctly assigned to their subspecies of origin.

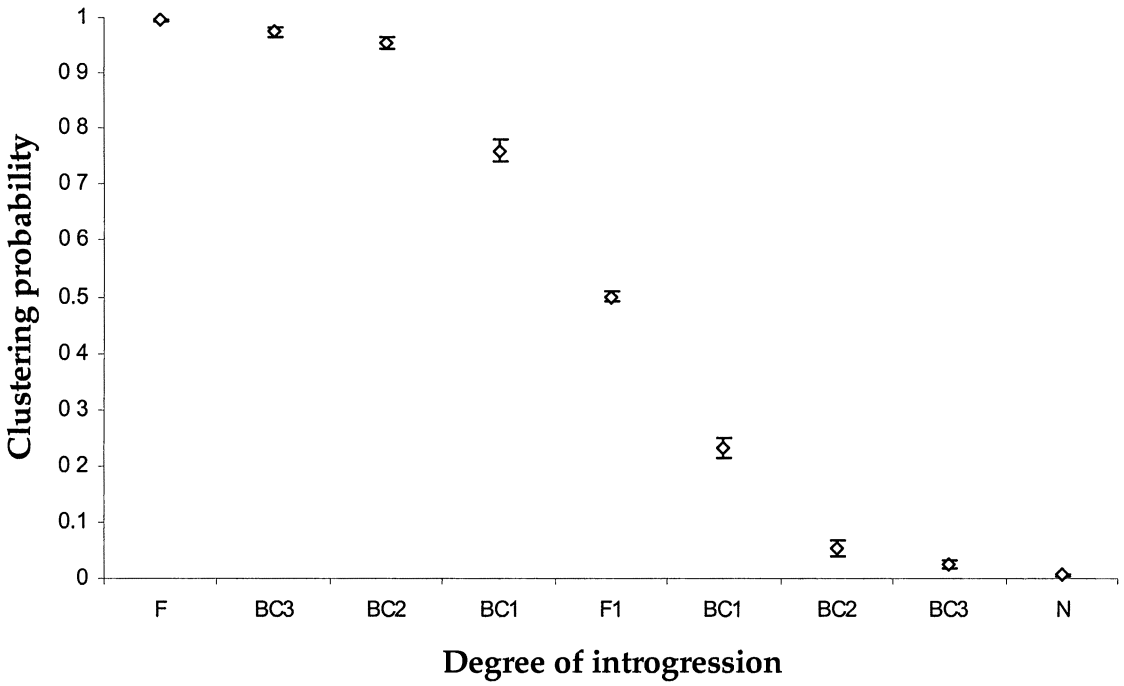

Figure 13. Bayesian determined clustering probabilities to cluster 1 based on genetic variation at eleven microsatellite loci at four levels of simulated introgression (N = 180 / level of introgression) and *M. salmoides* populations on a scale from 0 (*M. s. salmoides* = N) to 1 (*M. s. floridanus* = F) (mean ± SE). The intergrade populations consist of the F_1 generation and three generations of parental backcrosses (BC).

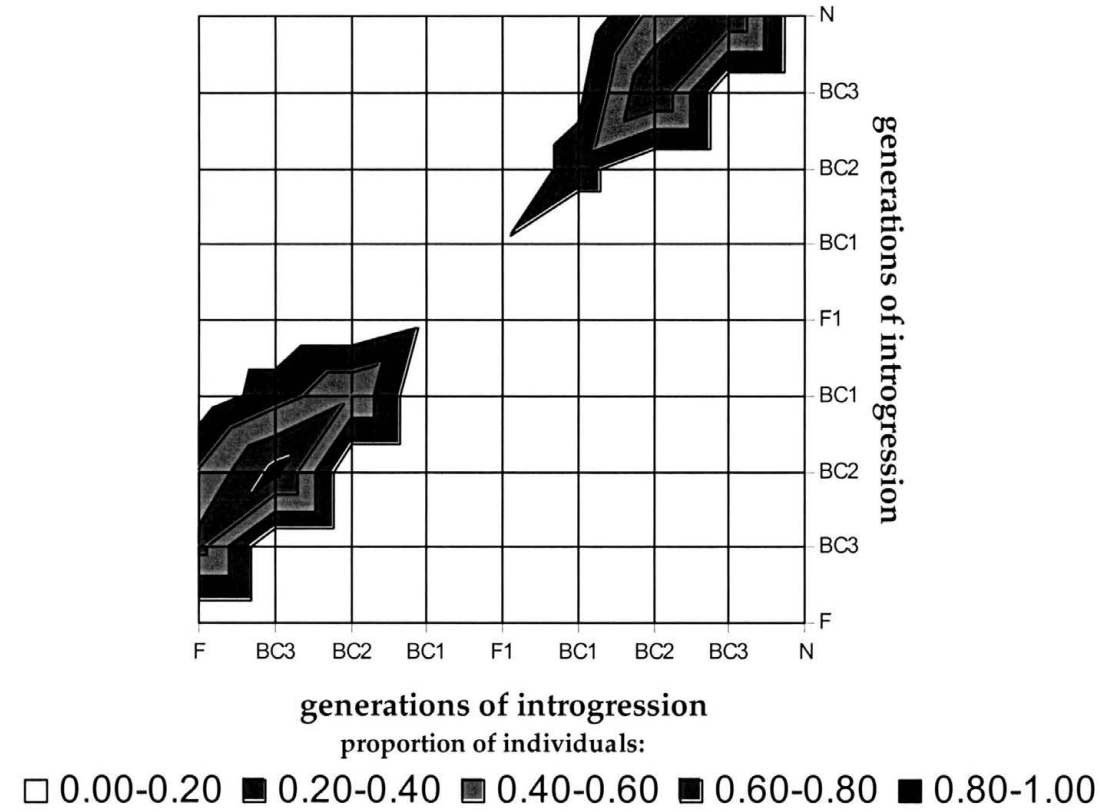

Figure 14. The proportion of individuals occurring within the range of clustering probabilities at more than one level of introgression based on the mean ± 3 σ of normalized data (*M. s. floridanus* = F; *M. s. salmoides* = N).

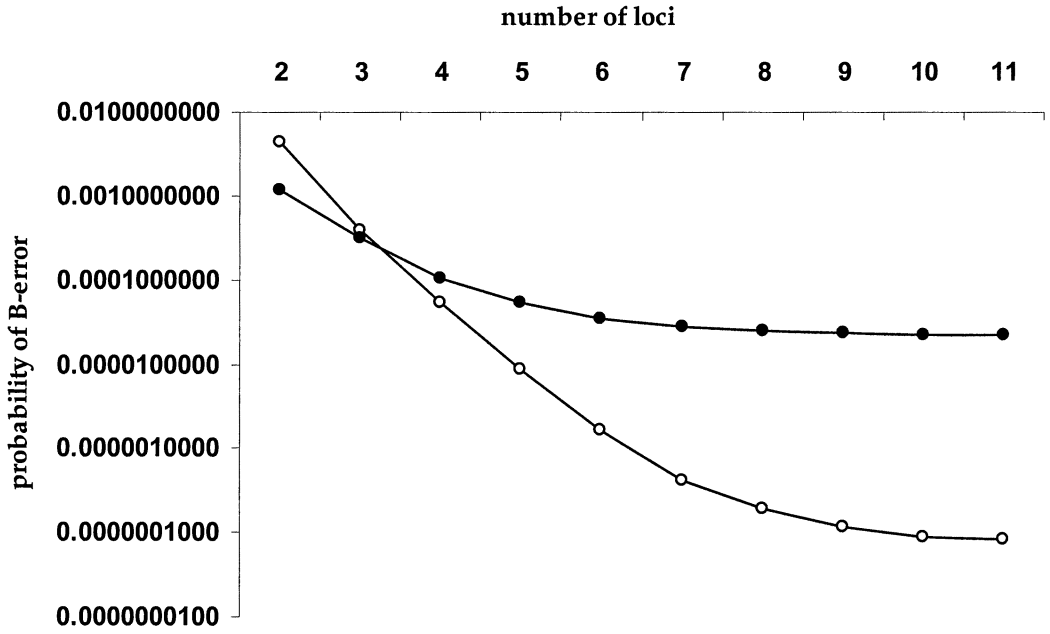

Figure 15. β-error probabilities (probability of duplicate genotypes) within each subspecies (assuming random mating among all individuals and populations), using locus combinations of two to eleven, from the most to least discriminatory. Black and white markers represent *M. s. floridanus* and *M. s. salmoides* respectively.

F_1 introgression:	NNNN **X** FFFF = N/F N/F N/F N/F
	⇩
Two Locus Combinations:	1.00(N/F N/F) = 1.00
	⇩
ID:	all correctly interpreted as F_1 generation individuals
	⇩
First Generation Backcross:	NFNF **X** NNNN = N/N F/N N/N F/N
	⇩
Two Locus Combinations:	0.67(N/N F/N or F/N N/N) + 0 17(N/N N/N) + 0.17(F/N F/N) = 1.00
	⇩
ID:	0.67 correctly inferred as first generation backcross individuals
	0.17 incorrectly inferred as northern subspecies
	0.17 incorrectly inferred as F_1 generation individuals

Figure 16. Proportion of incorrect assignments from sampling two loci in a 1st generation parental backcross (PBC), assuming each locus has two alleles fixed in alternate subspecies.

WORKS CITED

Allendorf, F. W. 1991. Ecological and genetic effects of fish introductions: synthesis and recommendations. Canadian Journal of Fisheries and Aquatic Science 48:17-181.

Angers, B., L. Bernatchez, A. Angers, and L. Desgroseillers. 1995. Specific microsatellite loci of brook charr (*Salvelinus fontinalis* Mitchill) reveals strong population subdivision on microgeographic scale. Journal of Fish Biology 47:177–185.

Arnaud, E., P. Presa, F. Krieg, D. Vaiman, and R. Guyomard. 1993. (CT)n and (GT)n microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). Heredity 71:488-496.

Avise, J. C. 1974. Systematic value of electrophoretic data. Systematic Zoology 23:465-481.

Bailey, R. M., and C. L. Hubbs. 1949. The black basses (*Micropterus*) of Florida, with description of a new species. University of Michigan Museum of Zoology Occasional Papers 516:1-40.

Beaty, P. R., and W. F. Childers. 1980. Introgression of northern largemouth bass *Micropterus salmoides salmoides* and northern smallmouth bass *Micropterus dolomieui dolomieui*. Illinois Natural History Survey, Institute of Natural Resources, Final Report to Bass Research Foundation, Starkville, Mississippi.

Bernatchez, L., and J. J. Dodson. 1991. Phylogenetic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. Evolution 45:1016-1035.

Bernatchez, L., and P. Duchesne. 2000. Individual-based genotype analysis in studies of parentage and population assignment : how many loci, how many alleles? Canadian Journal of Fisheries and Aquatic Science 57:1-12.

Bianchini, M. L. 1995. Species introductions in the aquatic environment: changes in biodiversity and economics of exploitation. Pages 213-222 *in* D. Philipp, J. M. Epifanio, J. E. Marsden, J. Claussen, and R. J. Wolotira, Jr., editors. Protection of Aquatic Biodiversity. Proceedings of the World Fisheries Congress, Vol. 3. Oxford and IBH Publishing, New Delhi.

Blouin, M. S., M. Parsons, V. Lacaille, and S. Lotz. 1996. Use of microsatellite loci to classify individuals by relatedness. Molecular Ecology 5:393-401.

Bottroff, L. J., and M. E. Lembeck. 1978. Fishery trends in reservoirs of San Diego County, California, following the introduction of Florida largemouth bass M. s. floridanus. California Department of Fish and Game 64:4-23.

Brookfield, J. F. Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5:453-455.

Buchanan, F. C., L. J. Adams, R. P. Littlejohn, J. F. Maddox, and A. M. Crawford. 1994. Determination of evolutionary relationships among sheep breeds using microsatellites. Genomics 22:397-403.

Cavalli-Sforza, L. L., and A. W. F. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 32:550-570.

Chakraborty R., M. De Andrade, S. P. Daiger, B. Budowle. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics 56:45-47.

Clugston, J. P. 1964. Growth of Florida largemouth bass, *Micropterus salmoides floridanus*, and northern largemouth bass, *M. s. salmoides*, in subtropical Florida. Transactions of the American Fisheries Society 93:146-154.

Colbourne, J. K., B. D. Neff, J. M. Wright, and M. R. Gross. 1996. DNA fingerprinting of bluegill sunfish Lepomis macrochirus using (GT)n microsatellites and its potential for assessment of mating success. Canadian Journal of Fisheries and Aquatic Science 53:342-349.

Cornuet, J-M., S. Piry, G. Luikart, A. Estoup, and M. Solignac. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153:1989-2000.

Crawford, A. M., S. M. Kappes, K. A. Paterson, M. J. deGotari, K. G. Dodds, B. A. Freking, R. T. Stone, and C. W. Beattie. 1998. Microsatellite evolution: testing the ascertainment bias hypothesis. Journal of Molecular Evolution 46:256-60.

DeWoody, J. A. and J. C. Avise. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. Journal of Fish Biology 56:461–473.

DeWoody, J. A., D. Fletcher, S. D. Wilkins, W. Nelson, and J. C. Avise. 2000. Molecular genetic dissection of spawning, parentage, and reproductive tactics in a population of redbreast sunfish, *Lepomis auritus*. Evolution 52:1802-1810.

Dunham, R. A., C. J. Turner and W. C. Reeves. 1992. Introgression of the Florida largemouth bass genome into native populations in Alabama public lakes. North American Journal of Fisheries Management 12:494-498.

Edwards, R. J. 1980. The ecology and geographic variation of the Guadalupe bass, *Micropterus treculi*. Doctoral dissertation. The University of Texas, Austin, Texas.

Efron, B. 1983. Estimating the error rate of a prediction rule: Improvements on cross-validation. Journal of the American Statistical Association 78:316-331.

Ellegren, H, C. R. Primmer, B. C. Sheldon. 1995. Microsatellite evolution: directionality or bias. Nat. Genetics 11:360-362.

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

Felsenstein, J. 1989. PHYLIP: Phylogeny Inference Package (version 3.2). Cladistics 5:164-166.

Fields, R., S. S. Lowe, C. Kaminski, G. Whitt, and D. P. Philipp. 1987. Critical and chronic thermal maxima of northern and Florida largemouth bass and their reciprocal F1 and F2 hybrids. Transactions of the American Fisheries Society 116:856-863.

Forshage, A. A., and L. T. Fries. 1995. Evaluation of the Florida largemouth bass in Texas, 1972-1993. American Fisheries Society Symposium 15:484-491.

Fries, L. T., G. L. Kurten, J, Isaac Jr., T. Engelhardt, D. Lyon, and D. G. Smith. 2003. Mass production of polyploidy Florida largemouth bass for stocking public waters in Texas. *In* D. P. Philipp and M. S. Ridgway, editors. Black Bass: Ecology, Conservation and Management.

Garza, J. C. and N. B. Freimer. 1996. Homoplasy for size at microsatellite loci in humans and chimpanzees. Genome Research 6:211-217.

Guo, S. and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hansen, M. M., and V. Loeschcke. 1994. Effects of releasing hatchery-reared brown trout to wild trout populations. Pages 273-289 *in*, V. Loeschcke, J. Tomiuk, and S.K. Jain, editors. Conservation genetics. Birkhauser Verlag, Basel.

Harvey, W. D., R. L. Noble, W. H. Neill, and J. E. Marks. 1984. A liver biopsy technique for electrophoretic evaluation of largemouth bass. Progressive Fish Culturalist 46:87-91.

Hines, S. A., D. P. Philipp, W. F. Childers and G. S. Whitt. 1983. Thermal kinetic differences between allelic isozymes of malate dehydrogenase (Mdh-B locus) of largemouth bass (*Micropterus salmoides*). Biochemical Genetics 21:1141-1151.

Hutchinson, W. F., G. R. Carvalho, and S. I. Rogers. 1999. A nondestructive technique for the recovery of DNA from dried fish otoliths for subsequent molecular genetic analysis. Molecular Ecology 8:893-894.

Inman, C. R., R. C. Dewey, and P. P. Durocher. 1978. Growth comparisons and catchability of three largemouth bass strains. Proceedings of the Annual Conference Southeastern Association of Game and Fish Commissioners 30:1-17.

Jeffreys, A. J, K. Tamaki,A. MacLeod, D. G. Monckton, D. L. Neil, and J. A. Armour. 1994. Complex gene conversion events in germline mutation at human minisatellites. Nat. Genetics 6:136-45.

Jin, L., Macaubast, C., Hallmayer, J., Kimura, A. and Mignot, E. 1996. Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. Proc. Natl. Acad. Sci. 93:15285–15288.

Johnson, D. L. 1975. A comparison of Florida and northern largemouth bass in Missouri. Doctoral dissertation. University of Missouri, Columbia, Missouri.

Johnson, R. L., and J. Pignature. 1995. Restriction fragment length polymorphisms in largemouth bass *Micropterus salmoides salmoides* in a small Massachusetts Kettlehole. American Midland Naturalist 133:364-367.

Jones, A. G., S. Ostlund-Nilsson and J. C. Avise. 1998. A microsatellite assessment of sneaked fertilizations and egg thievery in the fifteen spine stickleback. Evolution 52:548-858.

Kassler, T. W., J. B. Koppelman, T. J. Near, C. B. Dillman, J. M. Levengood, D. L. Swofford, J. L. VanOrman, J. E. Claussen, and D. P. Philipp. 2002. Molecular and morphological analyses of the black basses *Micropterus*: Implications for taxonomy and conservation. American Fisheries Society Symposium 31:291-322.

Kimura, M. and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725-738.

Kleinsasser, L. J., J. H. Williamson, and B. G. Whiteside. 1990. Growth and catchability of northern, Florida, and F1 hybrid largemouth bass in Texas ponds. North American Journal of Fisheries Management 10:462-468.

Landry, C. and L. Bernatchez. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatability complex and microsatellite loci in Atlantic salmon (*Salmo salar*). Molecular Ecology 10:2525-2539.

Lessa, E. P. 1990. Multidimensional analysis of geographic genetic structure. Systematic Zoology 39:242-252.

Lewin, R. 1989. Limits to DNA fingerprinting. Science 243:1549-1551.

Lewontin, R. C. and K. Kojima. 1960. The evolutionary dynamics of complex polymorphisms. Evolution 14:450-472.

Levinson, G. and G. A. Gutman. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Molecular and Biological Evolution 4:203-221.

MacCrimmon, H. R., and W. H. Robbins. 1975. Distribution of the black basses in North America. In: Black Bass Biology and Management, R.H. Stroud and H. Clepper, eds., Sport Fishing Institute, Washington, D. C.

Maceina, M. J., and B. R. Murphy. 1992. Stocking Florida largemouth bass outside of its native range. Transactions of the American Fisheries Society 121:686-691.

Malloy Jr., T. P., R. A. Van Den Bussche, W. D. Coughlin, and A. A. Echelle. 2000. Isolation and characterization of microsatellite loci in smallmouth bass *Micropterus dolomieu* (Teleosti: Centrarchidae), and cross-species amplification in spotted bass *M. punctulatus*. Molecular Ecology 9:1919-1952.

Michalakis, Y., and L. Excoffier. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142: 1061-1064.

Near, T. J., T. W. Kassler, J. P. Koppelman, K. B. Dillman, and D. P. Philipp. 2003. Speciation in North American black basses *Micropterus* (Actinopterygii: Centrarchidae). Evolution 57:1610-1621.

Neff, B. D., P. Fu, and M. R. Gross. 1999. Microsatellite evolution in sunfish (Centrarchidae). Canadian Journal of Fisheries and Aquatic Sciences 56:1198-1205.

Nielsen, E. E., M. M. Hansen, and V. Loeschcke. 1997. Analysis of microsatellite DNA from old scale samples of Atlantic salmon: a comparison of genetic composition over sixty years. Molecular Ecology 6:487-492.

Nielsen, E. E., M. M. Hansen, and V. Loeschcke. 1999a. Analysis of DNA from old scale samples: technical aspects, applications and perspectives for conservation. Hereditas 130:265-276.

Nielsen, E. E., M. M. Hansen, L. A. Bach. 2001. Looking for a needle in a haystack: Discovery of indigenous Atlantic salmon (*Salmo salar* L.) in stocked populations. Conservation Genetics 2: 219-232.

Ohta, T. and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetic Research 22:201-204.

O'Reilly, P. T., L. C. Hamilton, S. K. McConnell, and J. M. Wright. 1996. Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Canadian Journal of Fisheries and Aquatic Science 53:2292–2298.

Paetkau, D., W. Calvert, I. Stirling, and C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. Molecular Ecology 4:347–354.

Paetkau, D., L.P. Waitis, P.L. Clarkson, L. Criaghead, and C. Stroebeck. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genetics 147:1943–1957.

Perez-Enriquez, R., M. Takagi and N. Taniguchi. 1999. Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers. Aquaculture 173: 413–423.

Philipp, D. P. 1991. Genetic implications of introducing Florida largemouth bass *Micropterus salmoides floridanus*. Canadian Journal of Fisheries and Aquatic Sciences 48:58-65.

Philipp, D. P. 1992. Stocking Florida largemouth bass outside its native range. Transactions of the American Fisheries Society 121:686-691.

Philipp, D. P., and G. S. Whitt. 1991. Survival and growth of northern, Florida, and reciprocal F1 hybrid largemouth bass in central Illinois. Transactions of the American Fisheries Society 120:58-64.

Philipp, D. P., W. F. Childers, and G. S. Whitt. 1983. A biochemical genetic evaluation of northern and Florida subspecies of largemouth bass. Transactions of the American Fisheries Society 120:58-64.

Philipp, D. P., W. F. Childers, and G. S. Whitt. 1985. Correlations of allele frequencies with physical and environmental variables for populations of largemouth bass, *Micropterus salmoides* (Lacepede). Journal of Fish Biology 27:347-365.

Potter, B.A., and B.A. Barton. 1986. Stocking goals and criteria for restoration and enhancement of cold-water fisheries. R.H. Stroud, editor. In Fish culture in fisheries management. Amer. Fish. Soc. Maryland.

Primmer, C. R., M. T. Koskinen, M. T. and J. Piironen. 2000. The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. Proceedings of the Royal Society of London 267:1699 –1704.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population STRUCTURE using multilocus genotype data. Genetics 155:945-959.

Rannala, B. and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Acadamy of Science 94:9197-9201.

Raymond, M. and R. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86:248-249.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223-225.

Roques, S., P. Duchesne, and L. Bernatchez. 1999. Potential of microsatellites for individual assignment: the North Atlantic redfish (*Sebastes*) species complex as a case study. Molecular Ecology 8:1703-1717.

Rosenberg, N. A., T. Burke, K. Elo, M. W. Feldman, P. J. Freidlin, M. A. M. Groenen, J. Hillel, A. Maki-Tanila, M. Tixier-Boichard, A. Vignal, K. Wimmers, and S. Weigend. 2001. Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. Genetics 159:699-713.

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142:1357-1362.

Ruzzante, D. E. 1998. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. Canadian Journal of Fisheries and Aquatic Science 55:1-14.

Schneider, S., D. Roessli, and L. Excoffier. 2000. ARLEQUIN: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

Shafland, P. L. 1999. The introduced butterfly peacock (*Cichla ocellaris*) in Florida. *In* Fish Community Analyses. Reviews in Fisheries Science 7:71-94.

Shaw, P. W., C. Turan, J. M. Wright, M, O'Connell and G. R. Carvalho. 1999. Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. Heredity 83:490-499.

Sjogren, P. and P. Wyoni. 1994. Conservation genetics and detection of rare alleles in finite populations. Conservation Biology 8:267-270.

Slatkin, M. 1994. An exact test for neutrality based on the Ewens sampling distribution. Genetic Research 64:71-74.

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

Slatkin, M. and L. Excoffier. 1996. Testing for linkage disequilibrium in genotypic data using the EM algorithm. Heredity 76:377-383.

Smith, R. P., and J. L. Wilson. 1981. Growth and comparison of two subspecies of largemouth bass in Tennessee ponds. Proceedings of the Conference of the Southeastern Association of Fish and Wildlife Agencies 34:25-30.

Taylor, E. L, P. Vercoe, J. Cockrem, D. Groth, J. D. Wetherall, and G. B. Martin. 1999. Isolation and characterization of microsatellite loci in the emu, *Dromaius novaehollandiae*, and cross-species amplification within Ratitae. Molecular Ecology 8:1963-1964.

Taylor, E. L, P. Vercoe, J. Cockrem, D. Groth, J. D. Wetherall, and G. B. Martin. 1999. Isolation and characterization of microsatellite loci in the emu, *Dromaius novaehollandiae*, and cross-species amplification within Ratitae. Molecular Ecology 8:1963-1964.

Tautz, D., and M. Renz. 1984. Simple sequences are ubiquitous components of eukaryotic genomes. Nucleic Acids Research 12:4127-4138.

Tautz, D., M. Trick and G. A. Dover. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature 322:652-656.

Tibayrenc, M., K. Neubauer, C. Barnabé, F. Guerrini, D. Skarecky, and F. Ayala. 1993. Genetic characterization of six parasitic protozoa: parity between random-primer DNA typing and multilocus enzyme electrophoresis. Proceedings of the National Academy of Sciences 90:1335-1339.

United States (U. S.) Congress, Office of Technology Assessment. 1993. Harmful non-indigenous species in the United States. U.S. Government Printing Office OTA-F-565, Washington, DC. 391 pp.

United States Fish and Wildlife Service (USFWS). 2001. National survey of fishing, hunting, and wildlife-associated recreation.

United States Fish and Wildlife Service (USFWS). 1996. Black bass fishing in the U.S., addendum to the 1996 national survey of fishing, hunting, and wildlife-associated recreation.

Vasemagi, A. R. Gross, T. Paaver, M. Kangur, J. Nilsson, and L. O. Eriksson. 2001. Identification of the origin of an Atlantic salmon (*Salmo salar* L.) population in a recently recolonized river in the Baltic Sea. Molecular Ecology 10:2877-2882.

Waldbieser, G. C., and R. W. Wolters. 1999. Application of polymorphic microsatellite loci in a channel catfish (*Ictalurus punctatus*) breeding program. Journal of the World Aquaculture Society 30:256-262.

Waser, P. M. and C. Strobeck. 1998. Genetic signatures of interpopulation dispersal. Evolutionary Ecology13:43-44.

Was, A. and R. Wenne. 2002. Genetic differentiation in hatchery and wild sea trout (*Salmo trutta*) in the Southern Baltic at microsatellite loci. Aquaculture 204:493-506.

Weber, J. L., and May, P. E. 1989. Abundant class of human DNA polymorphisms which can be typed using polymerase chain reaction. American Journal of Human Genetics 44:388-396.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1359-1270.

Whitmore, D. H. and T. R. Heiller. 1988. Natural introgression between largemouth and smallmouth bass (*Micropterus*). Copeia 1988:672-679.

Williams, J. D., S. Kazianis, and R. B. Walter. 1998. Use of random amplified polymorphic DNA (RAPD) for identification of largemouth bass subspecies and their intergrades. Transactions of the American fisheries Society 127: 825-832.

Williamson, J. H., G. J. Carmicheal, M. E. Schmidt, and D. C. Morizot. 1986. New biochemical genetic markers for largemouth bass. Transactions of the American Fisheries Society 115:460-465.

Wright, J. M. 1993. DNA fingerprinting in fishes. In: Biochemistry and Molecular Biology of Fishes, Vol 2, Hochachka, P. W., and T. P. Mommsen, eds., Elsevier, Amsterdam.

Wright, G. L., and G. W. Wigtil. 1981. Comparison of growth, survival, and catchability of Florida, northern, and hybrid largemouth bass in a new Oklahoma reservoir. Proceedings of the Annual Conference of the Southeastern Association of Fish and Game Commissioners 34:31-38.

Yardley, D., J. C. Avise, J. W. Gibbons, and M. H. Smith. 1974. Biochemical genetics of sunfish, III. Genetic subdivisions of fish populations inhabiting heated waters from nuclear reactors. U. S. Atomic Energy Commission Symposium Series CONF-730505:255-263.

Zolcynski, S. J. Jr., and W. D. Davies. 1976. Growth characteristics of the northern and Florida subspecies of largemouth bass and their hybrid, and a comparison of catchability between the subspecies. Transactions of the American Fisheries Society 105:240-243.

VITA

Dijar Jaime Lutz-Carrillo was born July 30th, 1979 in Austin, Texas, the son of James T. Lutz and Alexandra Carrillo. After completing his work at Manor High School, Manor, Texas, in 1997, he attended the University of Texas-Austin, where he earned a Bachelor of Science degree in Marine and Freshwater Biology in 2001. Upon graduation he entered the Aquatic Biology graduate program of Texas State University-San Marcos. While in the graduate program he was employed with Texas Parks and Wildlife, as a Fish and Wildlife Technician, and taught laboratory courses in statistical analyses and modern biology.

Permenant Address: 13304 Dearbonne Drive

Del Valle, Texas 78617

This thesis was typed by Dijar J. Lutz-Carrillo.