Using genetic tools to understand the population ecology of stream fishes

James Henry Roberts III

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Fish and Wildlife Conservation

Paul L. Angermeier, Co-chairman
Eric M. Hallerman, Co-chairman
C. Andrew Dolloff
J. Paul Grobler
Marcella J. Kelly

April 27th, 2012

Blacksburg, Virginia

Keywords: darter, dispersal, endangered species, genetic diversity, landscape ecology, isolation-

by-distance, metapopulation dynamics, persistence, population structure, stream fish

Using genetic tools to understand the population ecology of stream fishes

James Henry Roberts III

ABSTRACT

Stream fishes are highly diverse, yet highly imperiled by human alterations of stream environments. Many species are poorly characterized with regard to the size and structure of populations and patterns of dispersal between populations, which complicates assessment of how human activities, both harmful and beneficial, will affect persistence. I used genetic tools to further this understanding in three case-study fish species of the southeastern United States: Roanoke logperch (*Percina rex*) of the greater Roanoke River basin and redline (*Etheostoma rufilineatum*) and greenside darters (*E. blennioides*) of the upper Tennessee River basin.

I found that endangered *P. rex* persists in seven isolated populations. Within populations, individuals exhibit extensive dispersal and gene flow, which maintains connectivity throughout entire watersheds. Most populations exhibit small contemporary effective population sizes and occupy few stream channels, and thereby face an elevated risk of extinction. Genetic estimates of divergence indicate that fragmentation was recent, coincident with the construction of major dams throughout the species' range. Close evolutionary relationships between most populations suggest that a translocation strategy could decrease extinction risks. I developed a framework to help guide the process of balancing small-population versus translocation risks when formulating conservation strategies. When the framework was applied to populations of *P. rex*, straightforward management prescriptions emerged. The framework also may prove useful for other fragmented species.

Unlike *P. rex*, *E. rufilineatum* and *E. blennioides* are relatively abundant where they occur. However, both species were strongly affected by fragmentation due to hydroelectric dams and reservoirs. Populations in small streams flowing directly into a reservoir had lower genetic diversity than populations in larger, more fluvially connected streams. Furthermore, indices of watershed urbanization (e.g., percent impervious surface, road density) were negatively correlated with genetic diversity and with a genetic index of population stability. This suggests that darters occupying isolated streams and/or urbanizing

watersheds experience smaller, more variable population sizes than darters elsewhere. Monitoring of

such genetic responses could provide a useful early indicator of ecosystem stress and a useful

complement to other biomonitoring techniques

ACKNOWLEDGMENTS

A PhD is a long journey that no one makes alone. My journey took longer than most, which enabled me to cross paths with many kindred spirits along the way. Over such a long period, some emotional ups and downs are inevitable. It was the support and fellowship of my personal and professional network that kept me progressing toward my goals.

My graduate committee showed respect for my ever-evolving ideas and patience as those ideas ran their course. Committee Co-Chair Paul Angermeier has borne with me for a particularly long time, from my Master's through my research position and now to the end of my PhD. I am forever grateful for his mentorship and friendship. My other Co-Chair Eric Hallerman, another good friend, has helped develop me into the molecular ecologist that I am. Committee members Andy Dolloff, Marcella Kelly and Paul Grobler have stuck with me through a project perched precariously at the intersection of several disciplinary frontiers – I thank them for their inquisitiveness and wisdom.

While at Tech I spent over a third of my life in the water, assisted by an army of venerable field technicians, volunteers, and fellow grad students. Special praise is due to Greg Anderson, Jane Argentina, Matt Bierlein, Jackie Brown, Amy Bush, Tye DeWebber, Corey Dunn, Dan Dutton, Jason Emmel, Mark Foster, Annie Grant, Dylan Hann, Pat Kroboth, Tim Lane, Michael Lang, Bryan Neuswanger, Brett Ostby, Phil Pegelow, Brandon Peoples, Jen Philhower-Gillen, Jeremy Pritt, Bjorn Schmidt, and Chris Williams, who put in extended details on Team Logperch. Baby Blood and I bid you thanks.

Once the field work was done and the fin clips were in the freezer, the tedium began. To that end, I was blessed to have the assistance of Dan Dutton and Joanne Printz, my lab-work gurus. They performed the vast majority of my extractions, PCR, and molecular troubleshooting, thereby preserving my sanity. I raise my pipette to you both.

The research the reader finds in this dissertation derived from a variety of small, cobbled-together projects that several agencies were kind enough to fund. I am thankful to have had the opportunity for collaboration with these agencies and the biologists that represent them. I particularly appreciate the

efforts of Mike Pinder and Scott Smith of VDGIF, William Hester and Kim Smith of USFWS, Phil Payonk and Michael Hosey of USACE, and Doug Beard of USGS to work with me to develop projects that met applied needs while enhancing the basic science of fish conservation.

I thank my friends near and far for keeping me centered over the years. I refrain from listing names to protect the innocent, but you know who you are.

Finally, my profoundest appreciation goes to my family – the Roberts women – who have shown infinite support and patience through an educational journey that seemed sometimes to have no end. My love for harlequined darters in sparkling streams is exceeded only by my love and admiration for Tonja, Holly, and Riley Roberts. I hope that they realize how much this dissertation owes to them.

We in southwest Virginia at the beginning of the 21st century are fortunate to live in a time and place where spectacular wildlife still swims in our backyards. This appreciation is not widely felt. Although this dissertation took a long time to write, it was but a millisecond of ecological time. I'm aware of no fish species that went extinct during the writing of this document, though some species undoubtedly moved farther down that path. As conservation biologists, we chronicle the status of the natural world via a series of snapshots. Some day my particular snapshots will seem quaint to the future reader. My hope, however, is that my photographic subjects still will be around, going about their mysterious lives on some gravelly shoal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

Stream fishes are among the most imperiled groups of North American fauna, due primarily to human alterations of streams and their watersheds (Ricciardi and Rasmussen 1999). Such alterations can decrease the quality and quantity of available habitat, which increases demographic and environmental stochasticity, inbreeding depression, and loss of adaptive potential, thereby accelerating rates of population extinction (Caughley 1994; Frankham et al. 2010). Moreover, aquatic ecosystems are particularly vulnerable to habitat fragmentation via the construction of anthropogenic barriers to dispersal (e.g., dams and roads; Fagan 2002; Morita and Yamamoto 2002). Isolation exacerbates the small-population threats described above by preventing demographic or genetic rescue, and may promote anthropogenically-driven genetic divergence of allopatric populations (Frankham et al. 2011).

Despite their conservation importance, most stream fishes are poorly characterized with regard to population structure, dynamics, and evolution (Schlosser and Angermeier 1995). The notable exception is salmonids, whose populations are commercially and culturally valuable and have been thoroughly investigated by ecologists and geneticists (reviewed in Rieman and Dunham 2000 and McElhany et al. 2000). However, theoretical and statistical models developed for salmonids may transfer poorly to other stream-fish taxa that exhibit contrasting life-histories and occupy different stream environments. For example, many Pacific salmonids of western North America migrate great distances over their lifetimes between freshwater spawning and marine growing habitats (Waples 1995). Despite this mobility, they exhibit strong fidelity to spawning habitats, which promotes significant genetic and life-history divergence over small spatial scales (Quinn et al. 2001). Other, "non-migratory" taxa such as North American darters, madtoms, and sculpin exhibit far less intra-specific life-history diversity and appear to complete their lifecycles within individual streams or watersheds (Roberts and Angermeier 2007). Differences in spatial ecology between salmonids and non-salmonids may result in different grains and extents of population structure, effective population sizes, and rates of inter-population dispersal. Better understanding of these attributes for non-salmonids would increase our general understanding of population processes in stream biota. Furthermore, studies of relationships between population processes

1

and anthropogenic modifications of the landscape (e.g., land-use conversion, dam construction) would permit more conservation-appropriate implementation and mitigation of such activities.

In light of these perceived needs, my dissertation has applied molecular genetic markers to better understand the ecology and evolution of three species of darters, and, where appropriate, generalize these findings to other stream fishes. Genetic markers provide novel tools for estimating demographic parameters, testing hypothesized models of metapopulation structure, and evaluating extinction risk of stream fishes (Koizumi et al. 2006; Lowe and Allendorf 2010). The overall goals of my dissertation were to improve scientific understanding of 1) the spatial grains and extents over which darter populations and metapopulations are organized, 2) how darter populations respond to anthropogenic modifications of the landscape, 3) how genetic and ecological data can be used to assess the history of and predict future risks to darter populations, and 4) how management guidelines can be prescribed based on this knowledge.

Questions pertaining to these goals were answered using data from two case studies. In the first, I focused on Roanoke logperch (*Percina rex*), an endangered species of darter endemic to a small number of streams in Virginia and North Carolina. Despite the species' conservation significance, prior to my study, relatively few data were available with which to assess the viability of logperch populations or their connectivity. I performed a range-wide genetic study of the species, characterizing genetic variation at 11 nuclear DNA microsatellite markers and one mitochondrial DNA (mtDNA) gene. I also collated ancillary data on the geographic size and juxtaposition of populations, relative density of *P. rex* within populations, and spatial distances, ecological differences, and degree of hydrologic fragmentation between populations. In Chapter 1, I examined range-wide patterns of genetic structure, estimated the total number and connectivity of *P. rex* populations, and evaluated whether natural or anthropogenic habitat boundaries best accounted for population boundaries. In Chapter 2, I narrowed my spatial focus and looked more intensively at ontogenetic patterns of dispersal and gene flow within the upper Roanoke River watershed. In Chapter 3, I developed a framework to assess relative risks to logperch populations incurred by the adoption of two alternative management strategies. I calculated risk metrics based on available genetic, demographic, and geographical data, and then applied risk-scoring criteria to metric

values to develop aggregate risk scores for two types of risk: "small-population risks" incurred by maintaining a small population in isolation versus "outbreeding-depression risks" incurred by translocating fish among genetically diverged populations. I discussed ways in which the framework could be modified for application to other fragmented populations, for example by changing the numerical risk thresholds, applying a weighting scheme to different metrics of risk, or incorporating uncertainty into risk scores.

The second case study involved two common species, redline (*Etheostoma rufilineatum*) and greenside darters (*E. blennioides*), in the upper Tennessee River basin of Virginia and Tennessee. For Chapter 4, I sampled populations of both species in streams throughout the basin, at sites that contrasted in patterns of stream size, degree of isolation, and land use. I then used a suite of microsatellite DNA markers to estimate the magnitudes of genetic diversity in darter populations and genetic differentiation between populations, presuming that these statistics served as proxies for demographic parameters such as population size and dispersal rate. Regression models were used to test which site characteristics, if any, best explained genetic variation among sites. I used results to discuss the potential for using population genetic data to monitor anthropogenic impacts to streams.

CHAPTER 1: Isolation by dams, not distance, creates discrete population structure in a riverine fish

ABSTRACT

As a general rule, gene flow between populations becomes less likely as spatial distance increases, which results in a pattern of genetic isolation-by-distance (IBD) among populations. However, the specific rate of IBD may vary considerably, depending on a species' intrinsic mobility, the juxtaposition of natural and anthropogenic barriers to migration, and the rate at which genetic drift (e.g., due to small population size) inflates genetic differentiation among locations. Such influences are poorly understood for many riverine species. I examined spatial patterns of population genetic structure and IBD in Roanoke logperch (*Percina rex*), a riverine fish, based on microsatellite DNA variation among individuals sampled from throughout the species' range. Multiple genetic clustering methods unambiguously delineated seven range-wide *P. rex* populations that exhibited strong differentiation from each other and no recent migrant exchange. The inferred grain of population structure most closely matched the watershed habitat scale, but population boundaries coincided more closely with hydroelectric dams than with natural habitat boundaries *per se*. Genetic differentiation between populations was weakly positively correlated with distance, but strongly negatively correlated with contemporary effective population size (N_e). Most populations exhibited small N_e and signatures of a recent bottleneck, suggesting that strong differentiation and weak regional IBD were recent phenomena. Within populations, I detected no subpopulation structure or IBD, suggesting panmixia maintained by extensive migration (up to 80 km). This information clarifies the importance of a watershed-grained perspective on conservation of *P. rex*. Overall, IBD models exhibited a poor fit to the observed genetic data at both the local extent (where migration overwhelmed drift) and the regional extent (where drift overwhelmed migration). Given the pervasiveness of anthropogenic fragmentation, such "island" population structures could be common among riverine biota. IBD, in contrast, may apply over only a narrow range of conditions that typically are unmet in contemporary riverine landscapes.

INTRODUCTION

Accurate delineation of population genetic structure is fundamental to understanding the demography and evolution of species. Expectations about structure impinge on all aspects of any genetic study. In the absence of data, such expectations typically are based on human perceptions of habitat structure and largely untested assumptions about the scales over which organisms view and respond to habitat. Rivers, for example, are ubiquitous landscapes, yet most of our understanding of the structure of riverine biota is based on a few taxa, such as salmonid fishes (e.g., Castric and Bernatchez 2004). Along with shorelines and other "linear" habitats, rivers are thought of as textbook examples of one-dimensional isolation-by-distance (IBD) environments (Hedrick 2009). In IBD models (e.g., Wright 1943; Kimura & Weiss 1964), mating occurs more frequently between adjacent individuals; at equilibrium, a positive relationship thus develops between the spatial and genetic distances separating pairs of individuals or subpopulations (Figure 1.1; Slatkin 1993; Rousset 1997; 2000; equivalent to Hutchison and Templeton's [1999] Case I).

Two lines of evidence support the applicability of IBD models to rivers. First, riverine habitat is patchy at multiple hierarchical scales (Frissell et al. 1986) and the distribution, abundance, movement, and structure of many riverine organisms responds to patch heterogeneity (Bunn & Hughes 1997; Matthews 1998). Thus, we might expect organisms to favor known, "home" habitat patches over distant, potentially unsuitable patches (Railsback et al. 1999). Second, fully aquatic riverine organisms exhibit a finite capacity for movement (Bunn & Hughes 1997; Rodriguez 2002) and must move within confined pathways (Fagan 2002). Such restricted movements might be expected to conform to the predictions of one-dimensional IBD models. Indeed, various tests for IBD in rivers have shown an overall positive relationship between genetic differentiation and the along-the-stream distance separating pairs of samples (e.g., Kelly & Rhymer 2005; Lowe et al. 2006; Whiteley et al. 2006).

Although positive IBD seems a reasonable expectation for riverine biota, alternative models also are plausible (e.g., Schlosser & Angermeier 1995; Tero et al. 2003). For example, extensive migration could overwhelm drift and produce a "flat" IBD pattern with low variance and weak differentiation

5

overall, equivalent to Hutchison and Templeton's (1999) Case II IBD pattern (Figure 1.1; Koizumi et al. 2006). This relationship could occur if migration distances are large relative to the spatial extent under consideration, migration is unaffected by patch boundaries, subpopulations are very large, or subpopulations were recently founded. Although movement distances of riverine species may be short on average, many exhibit leptokurtic movement distributions (Rodriguez 2002). The homogenizing effects of occasional long-distance dispersers might overwhelm differentiation at scales much larger than average movement distances would suggest and produce a Case-II IBD pattern at such scales. Such dynamics might be better characterized by an unstructured population model than by a structured model like IBD.

Conversely, strong drift could overwhelm migration and produce a "noisy" IBD pattern with high variance over all distances and strong differentiation overall, equivalent to Hutchison and Templeton's (1999) Case III IBD pattern (Figure 1.1; Tero et al. 2003; Koizumi et al. 2006). This relationship could occur if migration is hindered by patch boundaries or barriers, migration distances are small relative to inter-patch spacing, subpopulations are small or extinction-prone, or subpopulations have been isolated for a long time. Because movement pathways naturally are limited in river systems, rivers are easily fragmented by natural and anthropogenic barriers to migration (Warren & Pardew 1998; Lowe et al. 2006; Beneteau et al. 2009). Such fragmentation could produce Case-III IBD both by prohibiting the homogenizing effects of migration and by exacerbating the differentiating effects of drift, via the carving of formerly contiguous habitats into smaller remnant patches that maintain smaller effective population sizes. Rather than IBD, this scenario might be better characterized by an isolation (Nei & Chakravarti 1977) or metapopulation (Whitlock & McCauley 1990) model.

In practice, a species may exhibit any of these patterns of IBD at different spatial and temporal scales (e.g., Markwith & Scanlon 2007; Beneteau et al. 2009). One reasonable set of hypotheses is that: (a) migration >> drift (Case-II) at small spatial extents (i.e., within panmictic subpopulations), (b) the influence of migration relative to drift decreases with distance at intermediate extents (i.e., between subpopulations within a population), and (c) drift >> migration (Case III) at large spatial extents (i.e., between isolated populations; Figure 1.1). Examination of the locations at which transitions between

these cases occur (i.e., the dashed boxes in Figure 1.1) may provide insight into the spatial extent of migration, the influence of habitat patchiness or barriers on gene flow, and the resulting spatial grain of discrete subpopulation and population boundaries. Decomposing IBD into its Case-I, -II, and -III components may further reveal the spatial scales over which migration versus drift is the predominant evolutionary force (Hutchison and Templeton 1999; Koizumi et al. 2006), which could inform genetic restoration programs.

In this study, I examined the population genetic structure of Roanoke logperch (*Percina rex*), a riverine fish, at multiple hierarchical habitat scales. At each scale I asked: 1) Does discrete structure exist? 2) What landscape features (i.e., patch boundaries *versus* anthropogenic barriers) most closely correspond with population boundaries? 3) Is there evidence for IBD? and 4) What do patterns of structure and IBD tell us about the rates and spatial scaling of migration? Answers to these questions were used to assess conservation options for *P. rex* and draw inferences about the general applicability of IBD models to riverine biota.

METHODS

Study species and area

Percina rex is a large-bodied (to 165 mm total length) member of the darter subfamily (Percidae: Etheostomatinae), a speciose North American group of primarily stream-dwelling fishes (Jenkins and Burkhead 1994). *P. rex* is endemic to the Roanoke, Dan, and Nottoway river basins of Virginia and North Carolina, where it occupies small to medium sized rivers (Roberts and Rosenberger 2008; Figure 1.2). Occupied watersheds are separated by long, unoccupied and potentially uninhabitable stream-reaches, major hydroelectric projects, or both. These hydroelectric projects, completed between 1920 and 1964, might impede migration not only across the dam, but also through the upstream reservoir and through the hydrologically unstable tailrace downstream of the dam (e.g., Skalski et al. 2008). The Nottoway basin is further isolated from the Roanoke and Dan basins by brackish Albemarle Sound; dispersal of *P. rex* between the Nottoway and Roanoke rivers is presumed to have occurred only during historical stream capture events (Jenkins and Burkhead 1994).

7

Within occupied watersheds, *P. rex* are patchily distributed among microhabitats (i.e., 1-m^2 patches) and channel units (i.e., riffles, runs, and pools) lacking heavy silt deposition (Rosenberger & Angermeier 2003). The species exhibits iteroparity and overlapping generations, with an age-at-maturity of 2.5 years and a lifespan of 6.5 years; generation time is unknown (Jenkins & Burkhead 1994). Dispersal and migration patterns are not well understood, but several extensive (> 2-km) upstream and downstream movements opportunistically have been observed (Roberts et al. 2008). Owing to the species' specialized habitat requirements, limited extant range, and presumed decline, *P. rex* is listed as "Endangered" under the United States Endangered Species Act (U.S. Federal Register 54:34468-34472) and "Vulnerable" on the IUCN Red List (www.iucnredlist.org). Recovery goals focus on monitoring and increasing population sizes, ensuring evolutionary viability, and restoring population connectivity (Roberts & Rosenberger 2008).

Sample collection

Field collections of Roanoke logperch were made by me, the Virginia Department of Game and Inland Fisheries (VDGIF), and the North Carolina Wildlife Resources Commission between 2003 and 2008 throughout the species' known range. A small section of tissue was excised from the caudal fin of captured individuals, and then fish were returned alive to the stream. A total of 578 individual DNA samples were collected at a total of 35 sites (Table 1.1; Figure 1.2). Each site comprised an approximately 100-300-m-long reach of river. Geographic coordinates were obtained at the midpoint of each site from a handheld GPS receiver. Some sites were sampled in multiple years to estimate the magnitude of temporal genetic variation. Sample sizes varied among sites and years due to variance in our ability to find and capture *P. rex* (Table 1.1).

Laboratory methods

I extracted template DNA from whole tissue samples using a PureGene DNA Extraction Core Kit A (Gentra Systems, Minneapolis, Minnesota, USA). I genotyped samples at eleven microsatellite DNA loci (*Prex33*, *Prex36*, *Prex37*, *Prex38*, *Prex41*, *Prex42*, *Prex43*, *Prex44*, *Prex45*, *Prex46*, and *Prex47*) developed for *P. rex* by Dutton *et al.* (2008), using methods reported therein. Forward primers for these

8

loci were labeled using NED, VIC, PET or FAM fluorescent dye (Applied Biosystems, Inc., Foster City,

California, USA) and PCR was conducted in a MyCycler Thermal Cycler (BioRad, Hercules, California,

USA). Amplification products were separated in an ABI 3130 automated sequencer and sized in

GENEMAPPER 3.5 using a LIZ500HD size standard (Applied Biosystems, Inc., Foster City, California,

USA).

Data analysis

Lacking *a priori* knowledge of population structure, we tested for Hardy-Weinberg and linkage

equilibrium separately for each site in ARLEQUIN 3.11 (Excoffier *et al.* 2005). For these analyses, data

from different sampling years were pooled by site. Hardy-Weinberg tests employed 10^5 recorded Markov-

Chain-Monte-Carlo (MCMC) chains, following a burn-in of 10^3 chains, whereas linkage tests employed

10^5 randomizations. Test results were evaluated using a sequential Bonferroni adjustment for a global $\alpha =$

0.05.

I tested for discrete population structure using both individual- and group-centered methods.

Based on individual multilocus genotypes, the number of discrete genetic clusters (K) was estimated

using STRUCTURE 2.1 (Pritchard *et al.* 2000). I hypothesized that population structure might be

hierarchical, so I first estimated the number of populations (K_P) within the entire dataset, then estimated

the number of sub-populations (K_S) within each of the K_P inferred populations. For K_P, I evaluated K

values from 1 to 15, whereas for each K_S, I evaluated K values from 1 to x, where $x =$ one plus the number

of sites sampled in the population in question. All STRUCTURE models allowed for admixture and

correlation of allele frequencies among clusters and searched parameter space using 10^6 recorded MCMC

chains, following a burn-in of 10^5 chains. Five replicates were run for each K value, and the replicate with

the highest log-likelihood score was retained as the best estimate of the likelihood of that K value. I

compared the fit of alternative models (i.e., K values) using ΔAIC_c, which estimated how much less

information an alternative model contained than the highest-likelihood model [$AIC_{ci} = 2K_i - 2\log(L_i) +$

$((2K_i (K_i + 1))/(n_i - K_i - 1))$; and $\Delta AIC_{ci} = AIC_{ci} - AIC_{cmin}$ for model i with likelihood L and sample-size of

individuals n]. Models with $\Delta AIC_c < 2$ were considered as plausible as the best model (Burnham and Anderson 2004).

Once discrete population structure was delineated, I used STRUCTURE to estimate the probability that each individual was a first-generation immigrant from a population other than the one from which it was sampled. For this analysis, K was fixed at the optimal K_P determined in the previous analysis and capture location was used as a Bayesian prior. The model assumed a background migration rate of 0.05 and correlation of allele frequencies among populations and searched parameter space using 10^6 recorded MCMC chains, following a burn-in of 10^5 chains. I concluded that an individual was an immigrant if its probability of origination from another population was > 0.5.

I also investigated discrete population structure using two group-centered approaches. First, I conducted a UPGMA cluster analysis of sites based on Nei *et al.*'s (1983) genetic distance (D_A) in POPULATIONS 1.2.3 (O. Langella; http://bioinformatics.org/~tryphon/populations). Statistical support for topological splits was assessed by bootstrapping 10^4 times across loci. Second, I used an analysis of molecular variance (AMOVA) in ARLEQUIN to decompose total genetic variance into the following hierarchically-nested sources: major river basins (Roanoke, Dan, Nottoway), watersheds within basins (Roanoke, Pigg, Goose, Otter, lower Smith, middle Smith, upper Smith, Nottoway), streams within watersheds, sampling sites within streams, sampling years within sites, and residual variation among individuals within samples. The magnitude of variance at each hierarchical level was tested for equality with zero based on 10^4 random permutations of objects among groups at the level being tested.

Allele richness (A), unbiased gene diversity (H_E), and observed heterozygosity (H_O) were estimated for sites and populations in FSTAT 2.9.3 (J. Goudet; http://www2.unil.ch/popgen/softwares/fstat.htm). I estimated mean M across loci for each population in ARLEQUIN; M is the ratio of the number of alleles to the size-range of alleles within a population, an index that decreases following a population bottleneck (Garza and Williamson 2001). I asked whether M was lower than expected at demographic equilibrium by comparing each population's observed M to the 95% confidence interval of M based on 10^4 simulated equilibrium populations in Critical_M.exe (J.

Garza; http://swfsc.noaa.gov). Simulations required parameter estimates for the pre-bottleneck mutation-scaled effective population size (θ_P), the proportion of non-single-step microsatellite mutation events (p), and the mean size of non-single-step mutations (Δg). Lacking prior knowledge of these parameters, I conducted simulations under a range of plausible θ_P (2-20) and Δg (2.5-3.5) values. The value of p has less effect on M, so I used a single value (0.1) that appears typical of microsatellite loci in all simulations (Garza and Williamson 2001).

I investigated the potential influence of genetic drift on genetic structure and IBD by estimating the effective population size (N_e) of inferred populations. Various estimators of N_e from allele-frequency data are available, each with advantages and disadvantages (Wang 2005). I opted for a contemporary measure of N_e conveying information over ecological timescales and not requiring knowledge of marker mutation patterns. I estimated N_e for each population using the bias-corrected linkage disequilibrium method, as implemented in LDNe (Waples and Do 2008). Models assumed random mating and, to balance precision with accuracy, excluded alleles occurring at a frequency < 0.02 (Waples and Do 2010). I estimated 95% confidence intervals by jackknifing over loci. Because of their long lifespan and slowed growth after Age 1, it is difficult to determine the age of Age-2+ Roanoke logperch, which made up the majority of our samples. When samples are drawn from multiple cohorts in a species with overlapping generations, estimates of N_e are strictly equivalent neither to a true single-generation N_e nor to the effective number of breeders (N_b), but instead represent some intermediate quantity (Waples and Do 2010). Nevertheless, because this sampling phenomenon affected all populations in similar ways, I presume that inter-population variation in estimated N_e was proportional to relative variation in the true size of gene pools.

I tested for IBD at multiple spatial scales and biological domains. Estimates of genetic differentiation were calculated between pairs of populations ($n = 7$), sites ($n = 35$), and individuals ($n = 578$) in SPAGeDi 1.3 (Hardy and Vekemans 2002). For IBD analyses, pair-wise differentiation between sites and between populations was indexed using Weir and Cockerham's (1984) estimator of F_{ST}, whereas differentiation between individuals was estimated using Rousset's (2000) $\hat{a}$. I also used SPAGeDi to

calculate R_{ST}, an allele-size-based measure of population differentiation, and to compare R_{ST} to F_{ST} to determine whether stepwise mutation had contributed appreciably to population differentiation, indicative of long-term isolation (Hardy et al. 2003). Allele sizes were randomly permuted among allele states 10^4 times to test the null hypothesis that $R_{ST} = F_{ST}$. I estimated spatial distances between sites along stream channels in Google Earth 5.1 (http://www.google.com/earth). Mantel tests for significant association between matrices of pair-wise genetic differentiation and spatial distance were performed using 10^4 random permutations of matrix elements in the program ZT (Bonnet and Van de Peer 2002). At the population scale, I also tested the hypothesis that drift explained genetic differentiation better than did migration using a Mantel test for association between pair-wise genetic differentiation and the harmonic mean N_e of each population pair. I used partial Mantel tests in ZT to assess the effect of each factor (distance or N_e) on differentiation after accounting for the other factor. In all tests involving F_{ST}, the statistic was "linearized" [i.e., $F_{ST} / (1 - F_{ST})$] prior to analysis (Rousset 1997).

RESULTS

Hardy-Weinberg equilibrium was rejected in only 1 of 385 tests (locus *Prex44* in site RR10), indicating a lack of appreciable influence from null alleles or site-scale Wahlund effects. Linkage equilibrium was rejected in only 8 of 1925 tests, and 6 of these cases occurred at one site (BO2). I therefore retained data from all 11 loci for further analyses.

When all 578 individuals were considered, STRUCTURE results supported a model of genetic structure comprising seven discrete *P. rex* populations (Figure 1.3; Appendix A). Models with alternative K_P values had essentially no statistical support ($\Delta AIC_c > 10$; Burnham and Anderson 2004). Inferred population memberships corresponded closely with geographic sampling locations (Figure 1.2). The spatial grain of PGS most closely matched the watershed habitat grain, but population boundaries corresponded more closely with the distribution of hydroelectric projects than with natural habitat boundaries. The exception to this pattern was that individuals from Site LS1 clustered with individuals from sites LS2, LS3, TC1, and TC2 (collectively forming population LSMITH), despite the bisection of these sites by Martinsville dam. Most individuals shared the majority of their ancestry with other

12

individuals captured in the same population, although some admixture was evident between populations PIGG and GOOSE and between UROAN and USMITH. Data partitions at $K_P < 7$ were inconsistent with intuitive expectations based on the spatial juxtaposition of populations. For example, the best 3-cluster model grouped UROAN with geographically distant populations USMITH and LSMITH but split UROAN from proximate PIGG, GOOSE, and OTTER. Although regional PGS was well defined, STRUCTURE showed no evidence for discrete sub-population structure within any of the seven inferred populations (Appendix A). The optimal K_S value was one within all seven populations and models with alternative K_S values had essentially no statistical support ($\Delta AIC_c > 10$).

No individuals were inferred to be first-generation migrants between any of the seven populations delineated by STRUCTURE. All immigrant probabilities were well below our threshold of 0.5; the maximum probability observed was 0.13. Because the optimal K_S was one within all seven populations, I did not examine first-generation migration within populations.

Group-centered analyses of discrete population structure provided results concordant with those of individual-centered analyses. The UPGMA cluster analysis based on Nei *et al.*'s (1983) D_A grouped the 35 sampling sites into seven terminal clusters (Figure 1.4) that matched those of STRUCTURE results. Also like the STRUCTURE results, the cluster analysis indicated closer genetic relatedness of UROAN to geographically distant USMITH and LSMITH than to geographically proximate PIGG, GOOSE, and OTTER. In the AMOVA, the largest structural components of genetic variance were attributable to differences among the three major basins (11.0%) and among watersheds within basins (20.8%), whereas negligible variance was attributable to differences among streams within watersheds, sites within streams, or years within sites (< 0.5% combined; Table 1.2). Furthermore, as in previous analyses, there was weak support for higher-level structuring of populations at the basin scale; variation among watersheds within basins was greater than variation among basins. Although lower (downstream of Martinsville Dam) and middle (between Philpott and Martinsville dams) sections of Smith River were considered separate watersheds for AMOVA, high genetic similarity of these areas was indicated by STRUCTURE, the UPGMA tree, and F_{ST}, so I grouped them collectively as the "LSMITH" population for subsequent

analyses. Furthermore, due to the low temporal variance, I pooled years within sites for subsequent analyses.

Measures of genetic diversity and N_e varied widely among the seven inferred populations. Allele richness and gene diversity were highest in UROAN and USMITH and lowest in OTTER and GOOSE (Table 1.3). Mean contemporary N_e was negative for UROAN, indicating a very large population size not discernible from infinity (Waples and Do 2010); I therefore used the lower bound of the 95% confidence interval, 1781 individuals, as a minimum estimate of N_e for UROAN. Estimates of N_e from other populations ranged from 807 individuals in LSMITH to 61 individuals in GOOSE. Confidence limits of N_e were wide for all populations and most had no estimable upper bound.

Results of tests for population bottlenecks depended somewhat on the assumed microsatellite mutation model, indicating that five to seven populations had undergone a recent bottleneck (Table 1.3). Only UROAN exhibited an M value within the published range of Ms (0.823-0.926) from populations with demographically stable histories (Garza and Williamson 2001). Four populations (GOOSE, OTTER, LSMITH, and USMITH) exhibited Ms within or below the published range of Ms (0.599-0.693) from populations known to have gone through bottlenecks (Garza and Williamson 2001). The Ms of remaining populations (PIGG and NOTT) were intermediate to these two ranges. Under the most liberal model assumptions ($\theta_P = 2$, $\Delta g = 2.5$), all seven populations exhibited Ms below the 95% confidence interval of M in simulated equilibrium populations. However, most populations exhibited heterozygosities larger than expected at equilibrium for $\theta = 2$ (see equations 7.8c and 7.9b in Hedrick 2009), so tests with $\theta_P = 2$ may have been unduly liberal. Under a more conservative assumed $\theta_P = 20$, UROAN did not exhibit evidence for a bottleneck, whereas NOTT exhibited evidence for a bottleneck if Δg was 2.5 but not if Δg was 3.5. The remaining five populations demonstrated significant evidence for a bottleneck under all mutation models considered.

Analyses of IBD were conducted at the population, site, and individual levels. Based on permutation tests in SPAGeDi, I did not reject the hypothesis that $R_{ST} = F_{ST}$ overall or in any pair-wise comparison (all $P > 0.05$). I therefore concluded that stepwise mutation did not contribute appreciably to

population differentiation and focused further analyses on F_{ST} as the most precise measure of differentiation (Hardy et al. 2003). In comparisons among the seven inferred populations, global F_{ST} was 0.17. Estimates of F_{ST} between population pairs ranged widely from 0.04 to 0.33 (Figure 1.5), but all were significantly greater than zero ($P < 0.05$). Linearized pair-wise F_{ST} values were positively correlated with the waterway distance separating a population pair, but not significantly so ($r = 0.35$, $P = 0.171$; Appendix B). Lack of a strong trend and generally high differentiation and scatter over all distances were consistent with a Case-III IBD pattern (Figure 1.1). On the other hand, pairwise F_{ST} values were significantly negatively correlated with the mean N_e of the population pair ($r = -0.49$, $P = 0.004$; Appendix B). Partial Mantel tests produced outcomes identical to those of simple Mantel tests. Tests therefore indicated that drift overwhelmed migration at the among-population scale.

In comparisons among sites, waterway distance was not significantly related to pair-wise F_{ST} within any of the populations investigated (i.e., UROAN, PIGG, OTTER, LSMITH, NOTT; all $P > 0.05$; Figure 1.5; Appendix B). Site-level IBD could not be evaluated within GOOSE or USMITH, because only one pair of sites was sampled within each of these populations. Comparisons among individuals produced similar outcomes; waterway distance between individuals was not significantly related to genetic distance ($\hat{a}$) within any of the seven populations (all $P > 0.05$; Figure 1.5; Appendix B). Lack of a strong trend and generally weak differentiation over all distances were consistent with a Case-II IBD pattern. Tests therefore indicated that migration overwhelmed drift at the within-population scale.

DISCUSSION

Spatial scale and mechanisms of discrete population structure

Discrete population structure is a fundamental assumption of many evolutionary and demographic models (Hanski & Gilpin 1997; Waples & Gaggiotti 2006). This assumption has been questioned on the bases that: a) individuals do not necessarily aggregate into discernible colonies, b) movement distributions may be more continuous than discrete, c) demes may not be internally panmictic, and d) individuals are not ecologically interchangeable (see reviews by Guillot et al. 2009 and Hawkes 2009). In such cases, application of more spatially realistic and individual-based models may increase the

realism of genetic studies. This increased realism, however, comes at a cost to generality and utility: most management activities ultimately are aimed at populations, not individuals. Better understanding of when we can and cannot "scale up" to populations therefore serves both the theory and practice of molecular ecology.

My survey of genetic variation in *P. rex* unambiguously indicated the presence of seven discrete populations. Both the juxtaposition and discreteness of inferred population boundaries were corroborated by various individual- and group-centered analyses. The estimated degree of genetic differentiation among populations was large, admixture was low, and no recent migrants were inferred. Accordingly, I find no evidence for ongoing genetic or demographic exchange among these populations. Rather, they appear to be isolated and on independent demographic and evolutionary trajectories, a situation perhaps best characterized by an isolation (Nei & Chakravarti 1977) or nonequilibrium metapopulation (Schlosser & Angermeier 1995) model.

I tested for discrete population genetic structure at various hierarchical habitat scales presumed important to a riverine fish like *P. rex*, including stream-reaches, streams, connected watersheds, and large drainage basins (Frissell et al. 1986). The spatial grain of observed structure best matched the watershed habitat grain. In the AMOVA, most genetic variance was captured at the watershed grain, less variance was captured at larger grain sizes, and essentially no variance was captured at smaller grain sizes. Individual-based analyses in STRUCTURE corroborated the lack of discrete subpopulation structure at the grain sizes of streams or stream-reaches within watersheds. Despite *P. rex*'s discontinuous distribution among reach- and stream-scale habitats (Rosenberger & Angermeier 2003), this patchiness apparently did not lead to patchy subpopulation structure. This finding suggests that the migration of *P. rex* was relatively insensitive to reach- and stream-scale habitat boundaries. Evidence for influences of natural habitat boundaries on movements of other stream fishes is equivocal (Lonzarich et al. 2000; Gilliam & Fraser 2001; Roberts & Angermeier 2007). However, observed lack of subpopulation structure over local spatial scales contradicts conventional wisdom suggesting that stream fish populations are

regulated primarily by reach-scale factors and supports the adoption of a watershed-scale perspective on population regulation (Fausch et al. 2002).

Although *P. rex* populations apparently were organized at the grain of watersheds, population boundaries coincided more closely with anthropogenic barriers than with natural habitat boundaries *per se*. In particular, nearly all population boundaries coincided with a major hydroelectric dam. Dams are pervasive features of riverine landscapes, and with population growth, economic development, climate change, and water shortages predicted for the future, their prevalence is expected to increase (Postel 2000). Various studies have documented decreased gene flow among populations separated by dams and reservoirs and decreased genetic diversity within populations isolated by dams and reservoirs (Pritchard et al. 2007; Skalski et al. 2008; Beneteau et al. 2009). The adaptive consequences of this anthropogenically-induced PGS are only beginning to be investigated (Waples et al. 2007) and warrant additional research, particularly for non-anadromous species.

Although dams were major determinants of *P. rex* population boundaries, there were exceptions. Populations occupying the Roanoke and Dan basins clearly were differentiated, although not all are separated from each other by a dam (e.g., GOOSE and LSMITH). However, occupied areas are separated by inundated upper reaches of Kerr Reservoir, as well as a long (~280 km), presumably unoccupied river segment. Such long distances and low habitat quality may overwhelm *P. rex*'s capacity for dispersal. Similarly, GOOSE and OTTER were strongly differentiated (F_{ST} = 0.19) and inferred to have exchanged no recent migrants, although the mouths of these rivers enter the Roanoke River only 17 km from each other and are not separated by a dam. Apparently, hydrologically unstable conditions in the tailrace downstream of Leesville Reservoir make the Roanoke River unsuitable for use as a transit corridor. In contrast, the lower section of Smith River (Site LS1) was genetically indistinguishable from sites in the middle section (LS2, LS3, TC1, TC2), despite the bisection of these sections by Martinsville dam and reservoir. Although this dam is the oldest under consideration (completed in 1920), it also is the shortest (~10 m high) and impounds a relatively short reach of river (< 3 km). It therefore may permit enough

gene flow (even if one-way, *sensu* Whiteley et al. 2010) to homogenize the gene pools upstream and downstream of the dam.

Lack of regional IBD due to strong drift

I found no statistical evidence for positive linear IBD (Case I in Figure 1.1) over any spatial scale examined. Lack of Case-I IBD indicates a lack of migration-drift equilibrium, and dissection of IBD plots, for example by stratifying by disturbance history, colonization history, or spatial scale, can provide insight into the mechanisms contributing to disequilibrium (Hutchison and Templeton 1999). Spatial stratification of IBD plots for *P. rex* into within- versus between-watershed components revealed dramatic differences in the influences of migration and drift across spatial scales.

At the range-wide scale, distance was a poor predictor of genetic differentiation between *P. rex* populations. The deterministic influence of spatial distance on migration and IBD was mostly overwhelmed by non-spatial inter-population variation in drift. Such drift apparently was accelerated by low contemporary N_e due to recent bottlenecks in most populations. Unlike geographic distance, contemporary N_e was a good predictor of pair-wise differentiation between populations. Low *M* values in most *P. rex* populations were consistent with those observed in other species known to have gone through bottlenecks (Garza & Williamson 2001), suggesting that these populations experienced severe reductions in the past 10-200 generations (see Figures 4 & 5 of Garza & Williamson 2001). Detailed demographic histories for these populations are lacking, but populations occupying the Piedmont (i.e., PIGG, GOOSE, OTTER, USMITH, and LSMITH) may have been chronically impacted by anthropogenic sedimentation associated with poor land-use practices since the 1700s (Jenkins & Burkhead 1994). Furthermore, PIGG likely was bottlenecked by a chemical discharge in 1975 that killed all fish within a 36-km segment of the Pigg River. UROAN and NOTT, which did not exhibit strong evidence for bottlenecks, are geographically extensive populations that exhibit high local abundances (Rosenberger & Angermeier 2003), which may buffer against demographic fluctuations and local extinctions.

Contemporary migration among populations appears to be precluded by hydroelectric projects and poor habitat quality. Although habitat quality probably began to decline prior to the completion of

dams, several lines of evidence suggest that population fragmentation is a recent phenomenon. First, because most populations are small, the observed level of differentiation could have arisen over a relatively short time. For example, assuming panmixia between GOOSE and OTTER prior to completion of Leesville Reservoir in 1963, an N_e of 81 individuals (the harmonic mean contemporary N_e for these populations), a generation time of three years, and Nei & Chakravarti's (1977) isolation model, the observed F_{ST} (= 0.19) could have developed after only 34 generations, or 102 years, of isolation. If these two populations were not initially panmictic or if generation time is shorter, differentiation could have developed faster, potentially since Leesville Reservoir was built. Second, the only two populations that contain high levels of extant genetic diversity, UROAN and USMITH, are less differentiated from each other than from other, more geographically proximate populations. An identical pattern was found in a survey of mtDNA variation in *P. rex* (George et al. 2010). This suggests that UROAN and USMITH retain the signatures of historically higher gene flow between the Roanoke and Dan basins, whereas other populations have lost such signatures along with their genetic diversity and exhibit inflated differentiation due to contemporary drift. The high genetic diversity and large effective size of UROAN provides a key frame of reference for evaluating the current and potential historical diversity of other populations. Third, comparisons of F_{ST} to its analog R_{ST} indicated that stepwise mutation did not contribute significantly to differentiation, thereby suggesting that isolation is a recent phenomenon (Hardy et al. 2003). Thus, available evidence suggests that the small size and strong isolation of *P. rex* populations are anthropogenic in origin. Conservation strategies for *P. rex* therefore might focus on increasing the N_e of and gene flow among populations, through some combination of: a) increasing habitat area via habitat restoration, b) restoring river connectivity, and c) translocating fish among populations.

Lack of local IBD due to strong migration

I asked whether IBD existed among sample sites or individuals within populations. Given the geographic extensiveness of populations, I expected imperfect mixing and local IBD within watersheds. Fishes clearly recognize habitat boundaries at channel-unit and stream scales, as reflected in nonrandom patterns in species' distributions (Matthews 1998). Fishes often avoid crossing the unsuitable "matrix"

19

between suitable riffles and pools (Lonzarich et al. 2000; Roberts & Angermeier 2007) and can have difficulty crossing small natural and anthropogenic barriers (Schlosser 1995; Warren & Pardew 1998). Even within permeable environments, fishes typically exhibit spatially restricted movement due to a combination of territoriality, philopatry, limited swimming ability, and selection against dispersal into unknown habitats (Railsback et al. 1999; Rodriguez 2002). Moreover, a globally positive IBD trend is common in studies of riverine biota (e.g., Kelly & Rhymer 2005; Whiteley et al. 2006; Markwith & Scanlon 2007). However, in cases where IBD plots are decomposed into within- versus between-watershed components, a variety of local-scale IBD patterns are revealed, including Case-I (Koizumi et al. 2006; Primmer et al. 2006; Markwith & Scanlon 2007), Case-II (Whiteley et al. 2006; Waits et al. 2008; Beneteau et al. 2009), and Case-III (Tero et al. 2003; Castric & Bernatchez 2004; Koizumi et al. 2006) scenarios. Such variability implies that distance-mediated migration is not universal across all spatial scales and life-history types.

At the within-population scale, geographic distance was a poor predictor of genetic differentiation between pairs of *P. rex* sites or individuals. Differentiation was uniformly low over all spatial distances considered (up to 80 km), presumably because the homogenizing effects of migration overwhelmed the differentiating effects of drift. Case-II IBD manifested consistently across all populations despite probable inter-population differences in population density, demographic history, and habitat permeability (e.g., Koizumi et al. 2006). The frequency of migration necessary to maintain panmixia over such spatial extents is impressive, especially given the low estimated N_es of several populations. Although I cannot infer from these data whether such migration is a single- or multi-generation process, preliminary sibship reconstructions for this species suggest that even single cohorts undergo watershed-extent dispersal (J. Roberts, unpublished data; see also Danancher et al. 2008). My results suggest that monitoring and restoration efforts for *P. rex* should recognize the high within-watershed connectivity of this species.

Applicability of the IBD model to riverine biota

One- and two-dimensional formulations of IBD evolutionary models (e.g., Wright 1943; Kimura & Weiss 1964) are consistent with intuitive expectations about the migration of organisms in spatially

structured environments. Such models can be useful for predicting the spatial rate of divergence and spread of genes, estimating historical influences of migration versus drift, and estimating demographic parameters such as the neighborhood size (Guillot et al. 2009). Although IBD models assume migration-drift equilibrium among the set of samples being compared, violation of this assumption seems to have only minor effects on demographic estimation (Leblois et al. 2004) and can be leveraged to test hypotheses about demographic history (Hutchison & Templeton 1999; Koizumi et al. 2006). However, the effects of anthropogenically induced disequilibria on the applicability of IBD models and their ability to predict evolutionary patterns are poorly investigated in riverine landscapes.

Although rivers are viewed as textbook examples of one-dimensional IBD environments, results of our study and others suggest a lack of universal fit of IBD models to riverine biota. I emphasize this result not to set up IBD as a "straw man", but to suggest that other conceptualizations of distribution and migration might be more appropriate for riverine biota under certain conditions. In *P. rex*, for example, patterns of regional PGS were more consistent with an isolation (Nei & Chakravarti 1977) or nonequilibrium metapopulation (Whitlock & McCauley 1990; Schlosser & Angermeier 1995) model than an IBD model. Conversely, at the local scale, a structured population model like IBD (or more complicated models, e.g., Schlosser & Angermeier 1995; Tero et al. 2003) was unnecessary to explain the well-mixed distribution of genetic diversity within *P. rex* populations; each population could be considered a single evolutionary unit (*sensu* Waples & Gaggiotti 2006).

How transferable are these patterns to other riverine biota? I posit that the conditions necessary for positive IBD to develop, distance-mediated migration along with moderate drift, are relatively uncommon in contemporary riverine settings. At regional scales, contemporary migration often is precluded due to anthropogenic barriers related to dams, roads, and poor habitat quality (Jones et al. 2000; Morita & Yamamoto 2002). Although human transfers of fishes across drainage boundaries may somewhat counteract reduced gene flow, the spatial distribution of such transfers is unlikely to reflect natural patterns of migration (Rahel 2010). Meanwhile, genetic drift is inflated by human activities that reduce usable habitat area and decrease the stability of fish populations. I expect the net effect of

21

fragmentation and inflated drift to be a shift of the transition zone between Case I and Case III IBD toward smaller spatial extents (Figure 1.1).

At local scales, fragmentation is less likely, yet positive IBD will develop only if migration is attenuated by distance. In *P. rex*, migration was unrestricted enough to overwhelm drift over large spatial extents that were bounded only by anthropogenic barriers. Extensive migration shifts the transition zone between Case II and Case I IBD out farther in space than studies of fish movement would suggest (Rodriguez 2002; Figure 1.1). If the spatial *extent* of high migration approaches the spatial *grain* of habitat fragmentation, as appears to be the case for *P. rex*, positive IBD has insufficient room to manifest. Migration extent is well-studied for only a few riverine taxa, including: a) salmonid fishes, which exhibit strong philopatry (e.g., Castric & Bernatchez 2004; Primmer et al. 2006), b) aquatic plants, which exhibit passive dispersal by streamflow (e.g., Tero et al. 2003; Markwith & Scanlon 2007), and c) aquatic insects, many of which can disperse by flight as adults (Bunn & Hughes 1997; Wilcock et al. 2007). Other, less-studied groups may exhibit more or less extensive migration based on differences in life-history (e.g., Turner et al. 1996; Whiteley et al. 2006). However, direct and indirect estimates of movement suggest that watershed-scale dispersal is common across a variety of non-salmonid fishes (Gilliam & Fraser 2001; Albanese et al. 2004; Hitt & Angermeier 2008; Waits et al. 2008). Given the combination of extensive migration capability and pervasive anthropogenic fragmentation, discrete population structures may be common among contemporary riverine biota. In contrast, IBD may apply over only a narrow range of conditions that often are unmet in fragmented riverine landscapes.

auspices of Institutional Animal Care and Use Committee protocols 08-106FIW and 11-035-FIW at Virginia Tech.

<div align="center">REFERENCES</div>

Albanese, B.W., P.L. Angermeier, and S. Dorai-Raj. 2004. Ecological correlates of fish movement in a network of Virginia streams. Canadian Journal of Fisheries and Aquatic Sciences 61:857-869.

Beneteau, C.L., N.E. Mandrak, and D.D. Heath. 2009. The effects of river barriers and range expansion on the population genetic structure and stability in greenside darter (*Etheostoma blennioides*) populations. Conservation Genetics 10:477-487.

Bunn, S.E. and J.M. Hughes. 1997. Dispersal and recruitment in streams: evidence from genetic studies. Journal of the North American Benthological Society 16:338-346.

Burnham, K.P. and D.R. Anderson. 2004. Multimodel inference: understanding *AIC* and *BIC* in model selection. Sociological Methods and Research 33:261-304.

Castric, V. and L. Bernatchez. 2004. Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase of genetic differences with distance. Molecular Ecology 13:1299-1312.

Bonnet, E. and Y. Van de Peer. 2002. ZT: a software tool for simple and partial Mantel tests. Journal of Statistical Software 7:1-12.

Danancher, D., J.I. Izquierdo, and E. Garcia-Vasquez. 2008. Microsatellite analysis of relatedness structure in young of the year of the endangered *Zingel asper* (Percidae) and implications for conservation. Freshwater Biology 53:546-557.

Dutton, D.J., J.H. Roberts, P.L. Angermeier, and E.M. Hallerman. 2008. Microsatellite markers for endangered Roanoke logperch (*Percina rex*), and their potential utility for other species. Molecular Ecology Resources 8:831-834.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin version 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Fagan, W.F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. Ecology 83:3243-3249.

Frissell, C.A., W.J. Liss, C.E. Warren, and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10:199-214.

Garza, J.C. and E.G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. Molecular Ecology 10:305-318.

George, A.L., D.A. Neely, and R.L. Mayden. 2010. Comparative conservation genetics of two endangered darters, *Percina rex* and *Percina jenkinsi*. Southeastern Fishes Council Proceedings 52:1-12.

Gilliam, J.F. and D.F. Fraser. 2001. Movement in corridors: enhancement by predation threat, disturbance, and habitat structure. Ecology 82:258-273.

Fausch, K.D., C.E. Torgersen, C.V. Baxter, and H.W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52:483-498.

Guillot, G., R. Leblois, A. Coulon, and A.C. Frantz. 2009. Statistical methods in spatial genetics. Molecular Ecology 18:4734-4756.

Hanski, I.A. and M.E. Gilpin. 1997. Metapopulation Biology: Ecology, Genetics and Evolution, Academic Press, San Diego, California.

Hardy, O.J. and X. Vekemans. 2002. SPAGeDi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2:618-620.

Hardy, O.J., N. Charbonnel, H. Freville, and M. Heuertz. 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163:1467-1482.

Hawkes, C. 2009. Linking movement behavior, dispersal and population processes: is individual variation a key? Journal of Animal Ecology 78:894-906.

Hedrick, P.W. 2009. Genetics of Populations, 4th Edition. Jones and Bartlett Publishers, Sudbury, Massachusetts.

Hitt, N.P. and P.L. Angermeier. 2008. Evidence for fish dispersal from spatial analysis of stream network topology. Journal of the North American Benthological Society 27:304-320.

Hutchison, D.W. and A.R. Templeton. 1999. Correlation of pairwise genetic and geographic measures: inferring the relative influences of gene flow and drift on the distribution of genetic diversity. Evolution 53:1898-1914.

Jenkins, R.E. and N.M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Jones, J.A., F.J. Swanson, B.C. Wemple, and K.U. Snyder. 2000. Effects of roads on hydrology, geomorphology, and disturbance patches in stream networks. Conservation Biology 14:76-85.

Kelly, M.W. and J.M. Rhymer. 2005. Population genetic structure of a rare unionid (*Lampsilis cariosa*) in a recently glaciated landscape. Conservation Genetics 6:789-802.

Kimura, M. and G.H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561-576.

Koizumi, I., S. Yamamoto, and K. Maekawa. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden char. Molecular Ecology 15:3175-3189.

Leblois, R., F. Rousset, and A. Estoup. 2004. Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. Genetics 166:1081-1092.

Lonzarich, D.G., M.R. Lonzarich, and M.L. Warren. 2000. Effects of riffle length on the short-term movement of fishes among stream pools. Canadian Journal of Fisheries and Aquatic Sciences 57:1508-1514.

Lowe, W.H., G.E. Likens, M.A. McPeek, and D.C. Buso. 2006. Linking direct and indirect data on dispersal: isolation by slope in a headwater stream salamander. Ecology 87:334-339.

Markwith, S.H. and M.J. Scanlon. 2007. Multiscale analysis of *Hymenocallis coronaria* (Amaryllidaceae) genetic diversity, genetic structure, and gene movement under the influence of unidirectional gene flow. American Journal of Botany 94:151-160.

Matthews, W.E. 1998. Patterns in Freshwater Fish Ecology. Chapman and Hall, New York, New York.

Morita, K. and S. Yamamoto. 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. Conservation Biology 16:1318-1323.

Nei, M. and A. Chakravarti. 1977. Drift variances of F_{ST} and G_{ST} statistics obtained from a finite number of isolated populations. Theoretical Population Biology 11:307-325.

Nei, M., F. Tajima, and Y. Tateno. 1983. Accuracy of estimated phylogenetic trees from molecular data II: gene frequency data. Journal of Molecular Evolution 19:153-170.

Postel, S.L. 2000. Entering an era of water scarcity: the challenges ahead. Ecological Applications 10:941-948.

Primmer, C.R., A.J. Veselov, A. Zubchenko, A. Poututkin, I. Bakhmet, and M.T. Koskinen. 2006. Isolation by distance within a river system: genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. Molecular Ecology 15:653-666.

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

Pritchard, V.L., K. Jones, and D.E. Cowley. 2007. Genetic diversity within fragmented cutthroat trout populations. Transactions of the American Fisheries Society 136:606-623.

Rahel, F.J. 2010. Homogenization, differentiation, and the widespread alteration of fish faunas. American Fisheries Society Symposium 73:311-326.

Railsback, S.F., R.H. Lamberson, B.C. Harvey, and W.E. Duffy. 1999. Movement rules for individual-based models of stream fish. Ecological Modelling 123:73-89.

Roberts, J.H. and P.L. Angermeier. 2007. Spatiotemporal variability of stream habitat and movement of three species of fish. Oecologia 151:417-430.

Roberts, J.H. and A.E. Rosenberger. 2008. Threatened fishes of the world: *Percina rex* (Jordan and Evermann 1889) (Percidae). Environmental Biology of Fishes 83:439-440.

Roberts, J.H., A.E. Rosenberger, B.W. Albanese, and P.L. Angermeier. 2008. Movement patterns of endangered Roanoke logperch (*Percina rex*). Ecology of Freshwater Fish 17:374-381.

Rodriguez, M.A. 2002. Restricted movement in stream fish: the paradigm is incomplete, not lost. Ecology 83:1-13.

Rosenberger, A.E. and P.L. Angermeier. 2003. Ontogenetic shifts in habitat use by the endangered Roanoke logperch (*Percina rex*). Freshwater Biology 48:1563-1577.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145:1219-1228.

Rousset, F. 2000. Genetic differentiation between individuals. Journal of Evolutionary Biology 13:58-62.

Schlosser, I.J. 1995. Dispersal, boundary processes, and trophic-level interactions in streams adjacent to beaver ponds. Ecology 76:908-925.

Schlosser, I.J. and P.L. Angermeier. 1995. Spatial variation in demographic processes of lotic fishes: conceptual models, empirical evidence, and implications for conservation. American Fisheries Society Symposium 17:392-401.

Skalski, G.T., J.B. Landis, M.J. Grose, and S.P. Hudman. 2008. Genetic structure of creek chub, a headwater minnow, in an impounded river system. Transactions of the American Fisheries Society 137:962-975.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264-279.

Tero, N., J. Aspi, P. Siikamäki, A. Jäkäläniemi, and J. Tuomi. 2003. Genetic structure and gene flow in a metapopulation of an endangered plant species, *Silene tatarica*. Molecular Ecology 12:2073-2085.

Turner, T.F., J.C. Trexler, D.N. Kuhn, and H.W. Robison. 1996. Life-history variation and comparative phylogeography of darters (Pisces: Percidae) from the North American central highlands. Evolution 50:2023-2036.

Waits, E.R., M.J. Bagley, M.J. Blum, F.H. McCormick, and J.M. Lazorchak. 2008. Source-sink dynamics sustain central stonerollers (*Campostoma anomalum*) in a heavily urbanized catchment. Freshwater Biology 53:2061-2075.

Wang, J. 2005. Estimation of effective population sizes from data on genetic markers. Philosophical Transactions of the Royal Society of London B 360:1395-1409.

Waples, R.S. and C. Do. 2008. LDNe: A program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753-756.

Waples, R.S. and C. Do. 2010. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications 3:244-262.

Waples, R.S. and O. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419-1439.

Waples, R.S., R.W. Zabel, M.D. Scheuerell, and B.L. Sanderson. 2007. Evolutionary responses by native species to major anthropogenic changes to their ecosystems: Pacific salmon in the Columbia River hydropower system. Molecular Ecology 17:84-96.

Warren, M.L. and M.G. Pardew. 1998. Road crossings as barriers to small-stream fish movement. Transactions of the American Fisheries Society 127:637-644.

Weir, B.S. and C.C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.

Whiteley, A.R., P. Spruell, and F.W. Allendorf. 2006. Can common species provide valuable information for conservation? Molecular Ecology 15:2767-2786.

Whiteley, A.R., K. Hastings, J.K. Wenburg, C.A. Frissell, J.C. Martin, and F.W. Allendorf. 2010. Genetic variation and effective population size in isolated populations of coastal cutthroat trout. Conservation Genetics 11:1929-1943.

Whitlock, M.C. and D.E. McCauley. 1990. Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. Evolution 44:1717-1724.

Wilcock, H.R., M.W. Bruford, R.A. Nichols, and A.G. Hildrew. 2007. Landscape, habitat characteristics, and the genetic population structure of two caddisflies. Freshwater Biology 52:1907-1929.

Wright, S. 1943. Isolation by distance. Genetics 28:114-138.

Table 1.1. Locations, sample sizes (*n*), and genetic diversity estimates for 35 *Percina rex* sites sampled between 2003 and 2008. Sites are organized by major basin (italics), inferred population (bold), and stream. Allele richness (*A*), gene diversity (H_E), and observed heterozygosity (H_O) are estimated across all sampling years combined.

Location	Code	Latitude °N	Longitude °W	*n*	*A*	H_E	H_O
Roanoke basin							
upper Roanoke	**UROAN**						
Roanoke River	RR1	37.26	-79.91	13	8.2	0.840	0.817
	RR2	37.26	-79.94	18	9.3	0.845	0.837
	RR3	37.26	-79.96	24	9.7	0.832	0.782
	RR4	37.27	-79.96	16	8.5	0.822	0.811
	RR5	37.27	-79.98	12	7.6	0.808	0.784
	RR6	37.27	-80.01	12	8.0	0.843	0.818
	RR7	37.27	-80.02	10	7.4	0.836	0.764
	RR8	37.28	-80.05	27	9.4	0.837	0.818
	RR9	37.28	-80.06	15	8.5	0.834	0.824
	RR10	37.28	-80.09	16	8.9	0.846	0.789
	RR11	37.28	-80.11	11	7.2	0.803	0.835
	RR12	37.24	-80.20	8	7.0	0.838	0.727
North Fork Roanoke River	NF	37.21	-80.29	9	9.0	0.851	0.818
South Fork Roanoke River	SF	37.16	-80.25	19	7.3	0.834	0.831
Pigg	**PIGG**						
Pigg River	PR1	36.94	-79.77	70	6.0	0.653	0.642
	PR2	37.00	-79.86	13	4.4	0.610	0.564
Big Chestnut Creek	BC	36.91	-79.80	9	4.0	0.655	0.616
Goose	**GOOSE**						
Goose Creek	GC1	37.17	-79.52	6	3.1	0.532	0.448
	GC2	37.27	-79.59	28	3.5	0.528	0.528
Otter	**OTTER**						
Big Otter River	BO1	37.21	-79.30	14	3.0	0.533	0.513
	BO2	37.25	-79.35	36	3.7	0.568	0.598
	BO3	37.31	-79.39	13	2.8	0.536	0.573
	BO4	37.37	-79.42	26	3.5	0.569	0.549
Little Otter River	LO	37.28	-79.43	7	2.7	0.563	0.610

Table 1.1, continued

Dan basin

lower Smith　　　　　　　**LSMITH**

lower Smith River	LS1	36.50	-79.76	10	4.8	0.667	0.600
middle Smith River	LS2	36.71	-79.94	7	4.0	0.693	0.725
	LS3	36.72	-79.94	10	4.5	0.702	0.700
Town Creek	TC1	36.80	-80.00	11	4.7	0.637	0.620
	TC2	36.82	-80.00	9	4.5	0.683	0.636

upper Smith　　　　　　　**USMITH**

upper Smith River	US1	36.84	-80.15	37	8.5	0.793	0.776
	US2	36.81	-80.20	5	5.1	0.811	0.782

Chowan basin

Nottoway　　　　　　　**NOTT**

Nottoway River	NR1	36.85	-77.57	19	5.0	0.685	0.665
	NR2	36.90	-77.67	16	6.1	0.685	0.744
Stony Creek	SC1	36.97	-77.45	10	4.9	0.718	0.655
	SC2	37.06	-77.57	12	4.6	0.692	0.682

Table 1.2. Results of analysis of molecular variance among hierarchical habitat scales for *Percina rex*. Probability values are based on 10^4 random permutations of objects among groups at the level being tested.

Source	Number sampled	Molecular variance	Percentage of variance	*P*-value
Basins	3	1.295	11.0	0.0001
Watersheds within basins	7	2.448	20.8	0.0001
Streams within watersheds	14	0.047	0.4	0.9744
Reaches within streams	35	0.001	0.0	0.5130
Years within reaches	6	0.000	0.0	0.1337
Individuals within years	578	7.979	67.8	0.0001
Total		11.770	100.0	

Table 1.3. Estimates of genetic diversity for populations of *Percina rex*, including sample size (*n*), allele richness standardized to 34 individuals (A_{34}), gene diversity (H_E), observed heterozygosity (H_O), mean and 95% confidence limits of contemporary effective population size (N_e), and the ratio of allele number to allele size-range (*M*). Population abbreviations follow Table 1.1.

Population	*n*	A_{34}	H_E	H_O	N_e Mean (95% limits)	*M*
UROAN	210	9.9	0.835	0.807	∞ (1781, ∞)	0.848[a]
PIGG	92	5.3	0.650	0.628	672 (190, ∞)	0.720[c]
GOOSE	34	3.8	0.531	0.515	61 (22, ∞)	0.546[c]
OTTER	96	3.4	0.562	0.569	121 (51, 1775)	0.641[c]
LSMITH	47	5.4	0.680	0.650	807 (101, ∞)	0.643[c]
USMITH	42	8.0	0.794	0.777	143 (83, 408)	0.670[c]
NOTT	57	6.2	0.693	0.688	289 (114, ∞)	0.790[ab]

[a]Exhibited significant evidence for a bottleneck when assumed $\theta_P = 2$ and $\Delta g = 2.5$.
[b]Exhibited significant evidence for a bottleneck when assumed $\theta_P = 20$ and $\Delta g = 2.5$.
[c]Exhibited significant evidence for a bottleneck under all modeled parameter values.
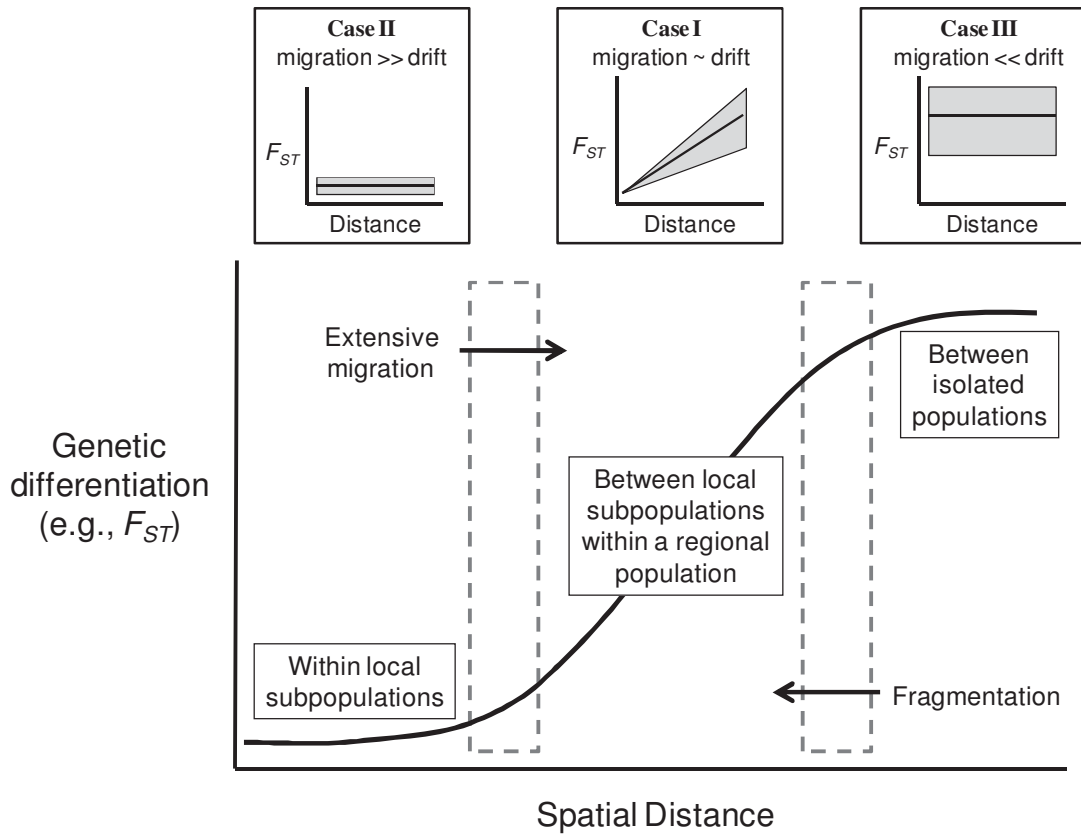
Figure 1.1. Theoretical predictions about relationships between genetic differentiation and spatial distance within and among populations. The expected mean and variance of relationships are represented by solid black lines and grey shaded areas, respectively. Arrows show how extensive migration and fragmentation shift the spatial locations of transition zones between cases (dashed boxes).
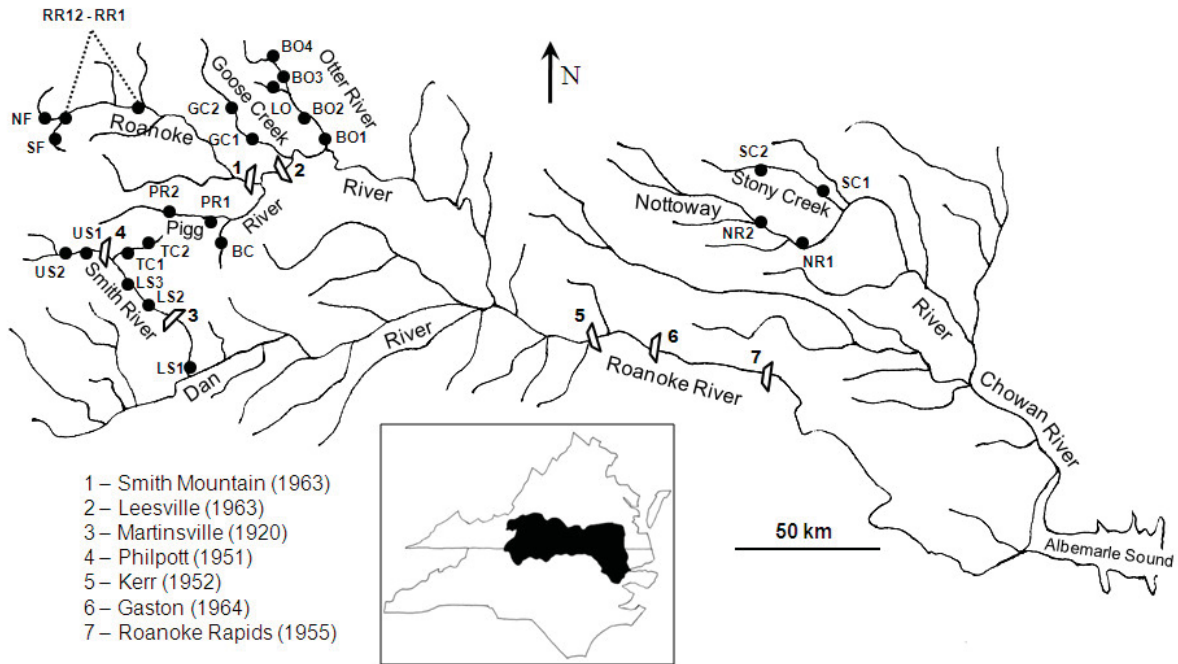
Figure 1.2. Locations of sites (filled circles) sampled for *Percina rex* within the Roanoke, Dan, and Nottoway basins of Virginia and North Carolina (see inset). Site names (in all capital letters) follow Table 1. Names and dates of completion of major hydroelectric dams (numbered trapezoids) are listed. For clarity, sites RR1-RR12 are not individually shown.
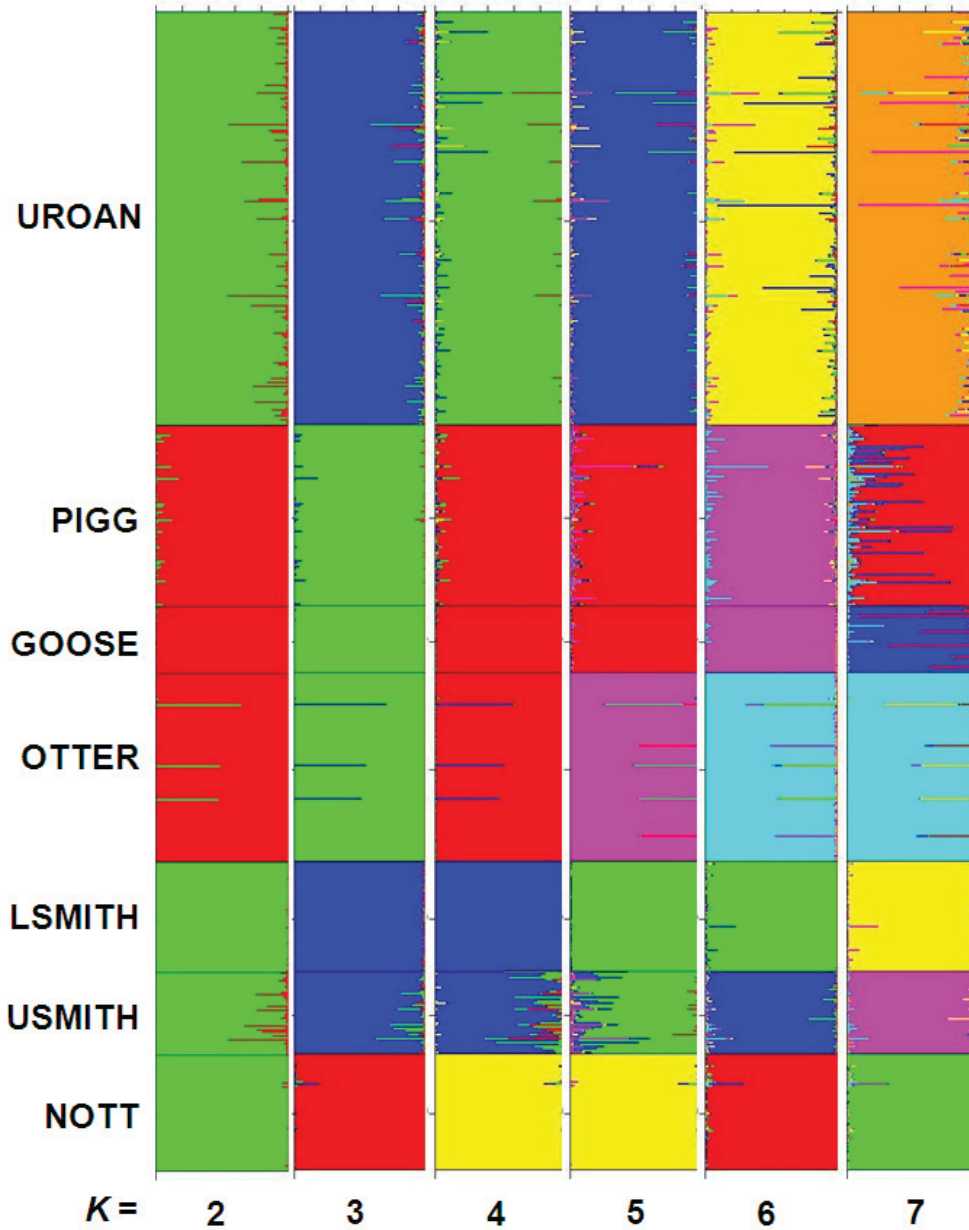
Figure 1.3. Comparison of STRUCTURE models with alternative hypothesized numbers of ancestral genetic clusters (*K*), given data from all 578 *Percina rex* individuals. Color coding indicates the proportion of each individual's ancestry (horizontal bars) originating from each of the *K* genetic clusters. The optimal model had seven clusters. Individuals are ordered by capture populations (delineated by thin black horizontal lines). Site codes follow those of Table 1.1.
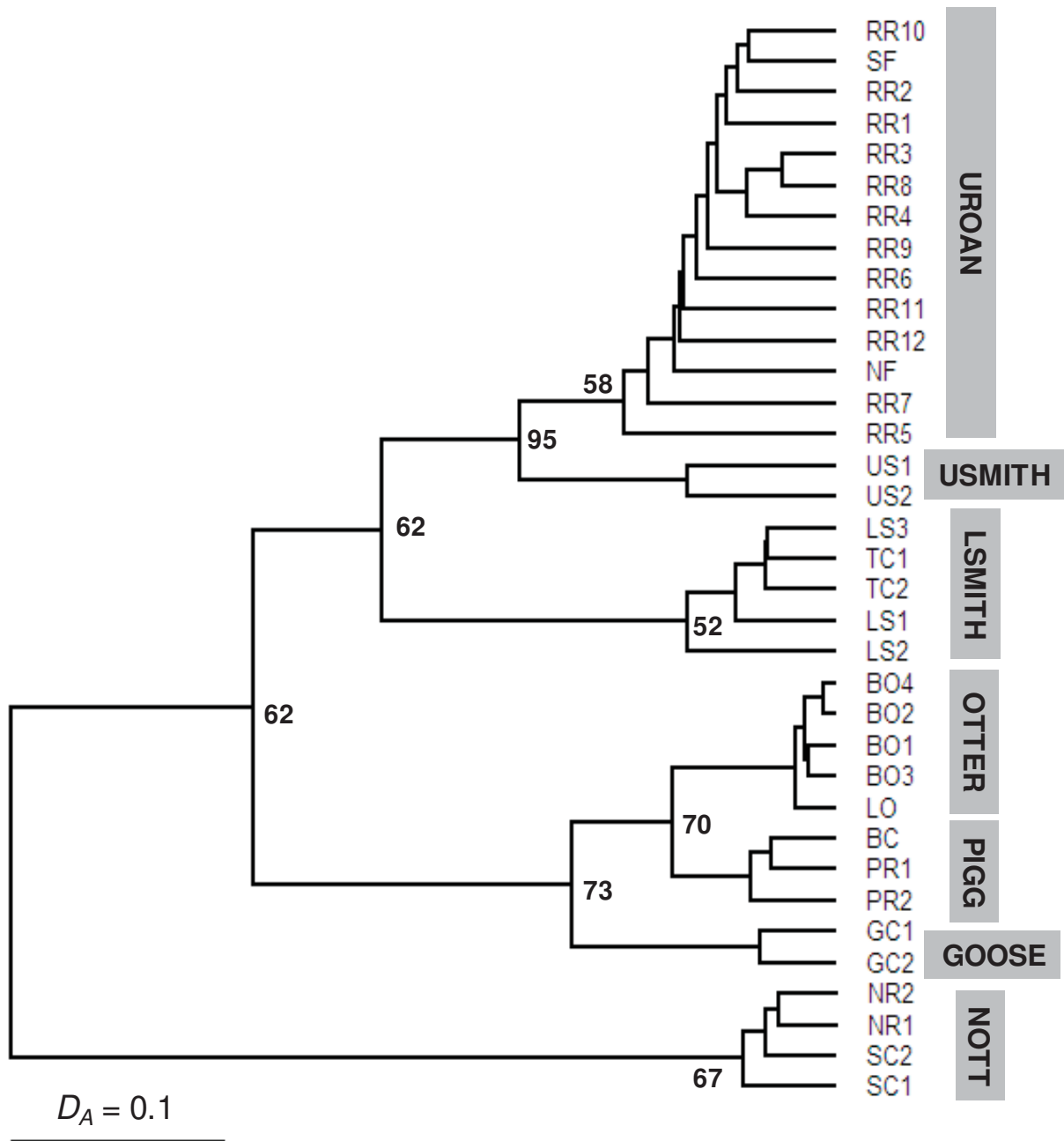
Figure 1.4. Results of a UPGMA cluster analysis based on Nei et al.'s (1983) genetic distance (DA) among 35 *Percina rex* sample sites. Probability values greater than 0.5, based on 104 bootstrap samples across loci, are shown for topological splits. Site codes follow those of Table 1.1.
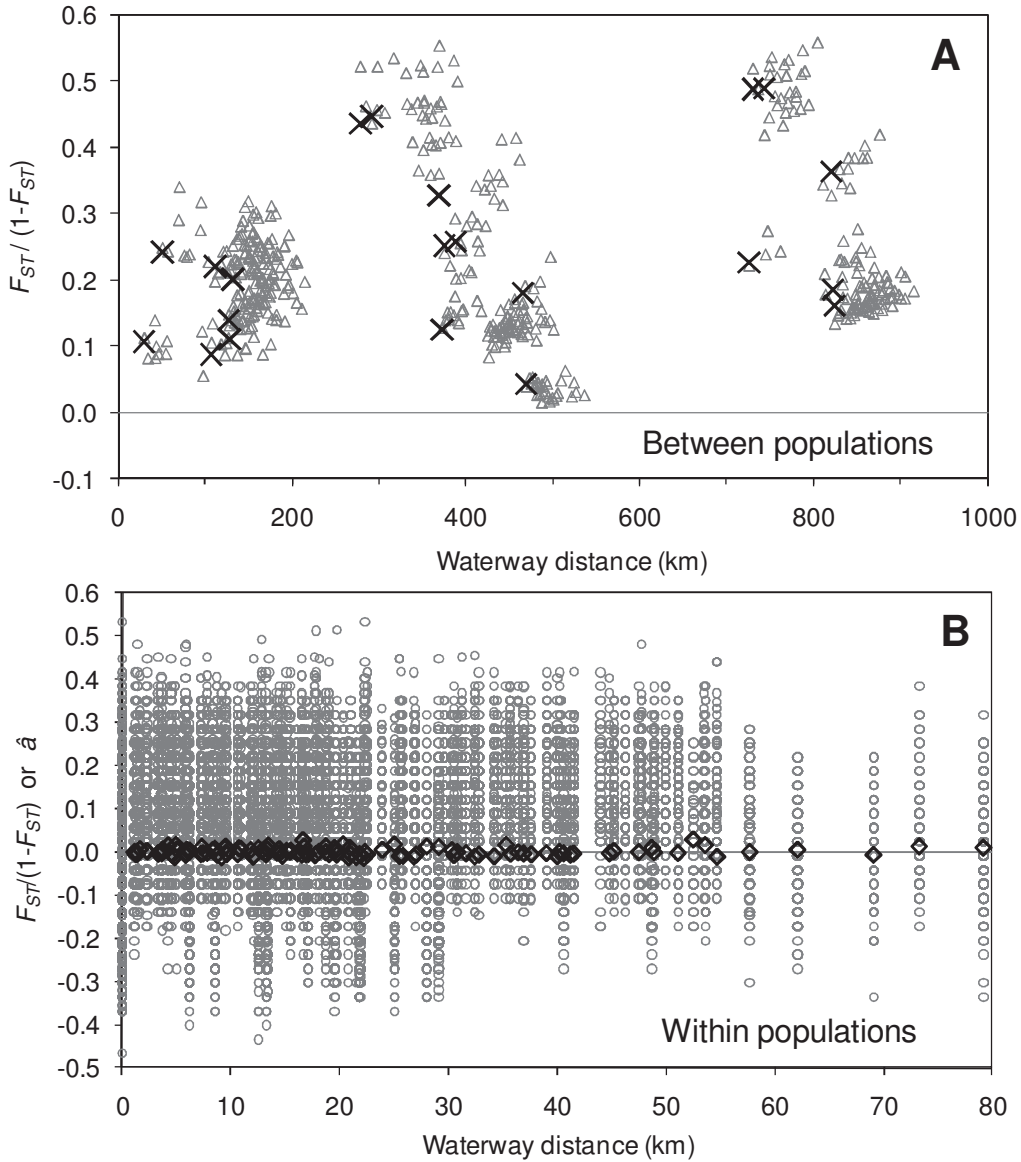
Figure 1.5. Relationships between genetic differentiation and waterway distance for pairs of *Percina rex* samples: A) Waterway distance between pairs of populations (black crosses) and pairs of sites located in different populations (grey triangles) is related to linearized F_{ST}. B) Waterway distance between pairs of sites located in the same population (black diamonds) and pairs of individuals located in the same population (grey circles) is related to linearized F_{ST} and $â$, respectively. None of these relationships exhibit slopes significantly different from zero (all $P > 0.05$; see Appendix B).

CHAPTER 2: Extensive dispersal of Roanoke logperch (*Percina rex*) inferred from genetic marker data

ABSTRACT

Most stream fishes are poorly characterized with respect to dispersal ecology, particularly during early life history. I used microsatellite DNA marker data to estimate dispersal and life-history parameters for a population of Roanoke logperch *Percina rex*, an endangered darter. Age-0 and Age-1 juveniles from the 2005 cohort ($n = 94$), as well as candidate parents ($n = 36$), were sampled for two years at sites throughout the upper Roanoke River watershed. Dispersal was inferred via three methods: 1) a genetic assignment test (AT) of individuals to their most likely site of origin, 2) spatial displacement of family members deduced through genetic pedigree reconstruction (PR), and 3) estimation of mean lifetime dispersal distance under an isolation-by-distance model. Reproductive success was widespread in 2005, as indicated by: a) large estimated effective number of breeders, b) rarity of closely-related individuals, and c) identical genetic diversity and weak genetic differentiation between juvenile and parental cohorts. Based on PR, polygamy was frequent among both sexes of parents, which spawned with an average of 2.4 mates. The sample contained 61 half-sibling pairs, but only one parent-offspring pair and no full sibs. Across methods, estimated dispersal of *P. rex* was extensive. The AT indicated unrestricted dispersal of juveniles among sites ≤ 15 km apart, while siblings inferred by the PR were captured an average of 14 km and up to 55 km apart. Extensive dispersion of siblings may have been accomplished by dispersal of parents between spawning events, post-spawn dispersal of juveniles, or both. No directional bias or temporal trend in dispersal was evident. Estimates of mean lifetime-dispersal distance (7-29 km, depending on method) bracketed the estimated average distance dispersed by juveniles, suggesting that a) the patterns I observed represented a typical cohort, b) the species performs most of its lifetime dispersal during the juvenile phase, and c) widely dispersed juveniles do, on average, successfully reproduce. Effective dispersal of *P. rex* was much more extensive than previous movement studies of darters would suggest, which indicates that monitoring and management activities for this population should target the entire watershed.

INTRODUCTION

Dispersal –the movement of individuals from one local breeding population to another – plays a key role in the fitness and persistence of stream fishes (Schlosser 1995a). To enhance fitness, individuals can "hedge bets" against environmental variability by spreading their reproductive output among locations (Winemiller and Rose 1992). Such dispersion could involve either the distribution of spawning effort across multiple habitats (i.e., "breeding dispersal") or the post-spawn movement of progeny away from a natal spawning habitat (i.e., "natal dispersal"; Greenwood 1980). In terms of population persistence, exchange of dispersers among local populations facilitates demographic supplementation and rescue (Schlosser 1995b; Labbe and Fausch 2000), re-colonization following local extirpations (Ensign et al. 1997; Gagen et al. 1998; Taylor and Warren 2001), and the maintenance of a longer persistence time and larger effective population size than a given local population could maintain in isolation (Hill et al. 2002; Hedrick 2009).

Despite the importance of dispersal to ecology and evolution, with the notable exception of salmonines (i.e., salmon, trout, and char), our understanding of the dispersal of stream fishes is limited. This limitation is due partly to a general sparseness of non-salmonine movement studies but mostly to methodological constraints that limit the application of existing movement data to questions about dispersal. Most stream fishes are too small-bodied to fit with a telemetry transmitter, which limits investigators to the use of capture-mark-recapture (CMR) techniques for these species or age-groups. Although CMR studies have advanced greatly in sophistication over the past decade (e.g., Rodriguez 2002; Labonne and Gaudin 2005), several disadvantages (relative to telemetry) remain, including: a) reliance on snapshot-in-time (versus continuous) data, b) low detection rates, c) high site-escapement rates, d) high intrusiveness during repeated recapture and handling (unless tags can be observed via passive means), and e) a constrained spatial and temporal scope. For example, CMR study areas in streams are typically < 10 km long (often much shorter) and study periods are typically ≤ 2 years (Fausch et al. 2002). Use of small study areas that are not closed to emigration injects well-known biases toward underestimation of movement distances (Gowan et al. 1994; Albanese et al. 2003; Schwalb et al. 2011).

Such biases undermine the characterization of dispersal, which likely occurs over larger spatiotemporal scales than are sampled by typical movement-study designs (Nathan et al. 2003). Another underappreciated limitation of traditional CMR studies is that the type of data they provide – temporal snapshots of the distance separating mark and recapture locations of fish – usually tells us little about the ultimate source (natal habitat) or destination (spawning habitat) of that fish. As a result, we cannot equate observed movements to an ecological function, such as habitat tracking (e.g., Fraser and Sise 1980), life-history expression (e.g., Northcote 1978), or dispersal (e.g., Quinn 1993).

Aforementioned limitations notwithstanding, salmonines may not be particularly good models for dispersal of other stream fishes. Many salmonines are highly migratory, spending their early life history in small streams but the adult portion of their lives in distant marine or lake habitats. Most other stream fish taxa do not make such dramatic habitat shifts over ontogeny, and therefore do not appear to migrate as extensively, though stream- to watershed-extent movements are common (Winn 1958; Hall 1972; Albanese et al. 2004; Hitt and Angermeier 2008). Furthermore, salmonines exhibit strong natal homing behavior (Northcote 1978), which also has been observed for some centrarchids (e.g., Ridgway and Shuter 1996) but not for many other stream fish taxa. Finally, migratory salmonines appear to perform the majority of their lifetime dispersal as adults (Quinn 1993; Hauser et al. 2006), whereas at least some centrarchids undergo greater dispersal as juveniles (e.g., Humston et al. 2009). The ontogeny of dispersal for many other taxa is completely unknown. Clearly, potential interspecific differences in life-history expression weaken the applicability of existing salmonine studies to questions about the dispersal of other taxa. New dispersal studies on non-salmonines would expand our general understanding of the spatial ecology of fishes in streams.

Genetic tools for examining stream-fish dispersal

The field of molecular population genetics provides three general techniques that complement CMR methods for understanding dispersal of stream fish: 1) indirect estimation of equilibrium dispersal or gene flow from theoretical models, 2) direct estimation of current dispersal using individual assignment tests, and 3) direct estimation of current dispersal based on family pedigree reconstruction.

Indirect dispersal estimation assumes demographic and genetic equilibrium over the recent evolutionary past, as well as some theoretical model that relates demography to genetics under an idealized population structure (e.g., Wright 1943; Takahata and Nei 1984). The equilibrium requirement can prove a major constraint (Whitlock and McCauley 1999), but the advantage of indirect methods is that they convey information about the ultimate consequence of dispersal: successful reproduction in a non-natal location, synthesized across many generations. Because dispersal may be infrequent and dispersing individuals difficult to detect in real time, this synthetic view of gene flow can complement direct study of dispersal in any given generation (Nathan et al. 2003). Although indirect methods are inappropriate for estimating gene flow from F_{ST} in an island model of migration (Whitlock and McCauley 1999), an alternative technique that estimates dispersal from genetic differentiation in an isolation-by-distance (IBD) model appears to be more robust to demographic disequilibria (Rousset 1997; 2000; Leblois et al. 2004) and better suited to linear stream environments. Despite its apparent utility, however, I am familiar with only one application of the IBD technique to the estimation of dispersal for a stream fish (i.e., Dolly varden char *Salvelinus malma*; Koizumi et al. 2006).

Unlike indirect equilibrium methods, assignment tests (ATs) can identify dispersing individuals in the current generation and provide a direct index of dispersal distance or rate (Manel et al. 2005). ATs use a discriminant function to assign each sampled individual to the group (i.e., population) from which it most likely originated. Assignments are based on the likelihood of obtaining an individual's genotype within a population, given that population's allele frequencies and the assumptions of Hardy-Weinberg and linkage equilibrium. Individuals assigned to a population other than the one in which they were sampled can be inferred to be immigrants from the former into the latter. Assignment methods have been used to detect dispersal (i.e., "straying") of various salmonine (e.g., Hauser et al. 2006; Lin et al. 2008) and non-salmonine (e.g., Hänfling and Weetman 2006; Beneteau et al. 2009; Raeymaekers et al. 2009) fishes. Although distinguishing ongoing dispersal from recent admixture can be difficult when populations are weakly differentiated (Waples and Gaggiotti 2006), ATs still are useful for examining spatial patterns in the relative magnitude of dispersal (Castric and Bernatchez 2004; Koizumi et al. 2006).

Like ATs, family pedigree reconstructions (PRs) enable the direct identification of dispersers in the current generation. Analysis of the post-spawn capture locations of parents, offspring, and siblings can be used to estimate the spatial dispersion of families over time (Danancher et al. 2008; Hudy et al. 2010; Kanno et al. 2011) and may reveal ontogenetic changes in dispersal behavior (Morrisey and Ferguson 2010; Vera et al. 2010). Although familial relationships typically are unknown, statistical methods allow the reconstruction of pedigrees from genetic marker data from a sample of individuals. A relatively new approach, which performs simultaneous group-likelihood assignment of parentage and sibship and allows for promiscuous or other mating systems, is particularly promising (Wang and Santure 2009). Because fishes commonly exhibit promiscuous mating systems (Helfman et al. 2009), PR approaches that allow for promiscuity should prove more accurate than those that do not. Furthermore, accounting for half-sib relationships provides an opportunity to separately examine the dispersion of full-versus half-sib pairs, which could provide insight into the prevalence of natal versus breeding dispersal, respectively.

Percina rex dispersal

I applied these genetic techniques to the dispersal ecology of Roanoke logperch (*Percina rex*), a benthic darter. *P. rex* is endemic to the Roanoke, Dan, and Nottoway river basins of Virginia and North Carolina, where it occupies rivers lacking heavy silt deposition (Rosenberger and Angermeier 2003; Roberts and Rosenberger 2008). Owing to the species' limited extant range and presumed decline, *P. rex* is listed as endangered under the U.S. Endangered Species Act (U.S. Federal Register 54:34468-34472). Recovery goals focus on monitoring and increasing population sizes, ensuring evolutionary viability, and restoring population connectivity. Achieving these goals will require an understanding of the species' spatial ecology, specifically: a) the distribution of spawning success across individuals and locations, b) the extent of dispersal and its variability over ontogeny, and c) the spatial scaling of population connectivity.

Better understanding of the dispersal ecology of *P. rex* will contribute not only to the conservation of this species, but also to our understanding and management of other darters and other

benthic fishes. Darters are a diverse, highly imperiled group for which relatively little large-scale

dispersal data exist. Although darters, like other benthic fishes, are often assumed to be sedentary, this

assumption is based on studies of limited spatial scope (Schwalb et al. 2010). Knowing the spatial habitat

needs of benthic fishes is important, because these needs determine how habitat fragmentation affects

population persistence of such species and dictates the spatial scales over which conservation

management should be directed.

Up to now, our understanding of the spatial ecology of *P. rex* primarily has been based on CMR

data (Roberts et al. 2008) and recently completed population genetic studies using microsatellite DNA

markers (see Chapter 1). Roberts et al. (2008) tracked movement of *P. rex* at 12 CMR sites in the upper

Roanoke River of Virginia over a 9-year period. Only 2 of the 22 marked *P. rex* that were recaptured,

both juveniles at initial capture, exhibited between-site movements; one moved 3.2 km upstream over five

years and the other moved 2.5 km upstream over two years. However, the study primarily was designed

to detect within-site movements and likely underestimated the frequency and extent of between-site

dispersal. Interestingly, genetic data from seven *P. rex* populations across the species' range showed no

evidence for population structure over large spatial extents ($\leq$ 80 km), implying high watershed-scale

gene flow (see Chapter 1). However, it is unclear from these data whether extensive gene flow was due

to extensive dispersal events by single cohorts or whether it took multiple generations of "stepping stone"

gene flow to achieve watershed connectivity. Further, data on gene flow could not reveal whether

dispersal was performed primarily by juveniles or adults.

In the present study, I attempt to bridge the information gap between the information provided by

small-scale CMR studies and that provided by large-scale studies of gene flow. I use direct and indirect

genetic techniques to estimate spatiotemporal patterns of dispersal for a cohort of juvenile *P. rex* in the

upper Roanoke River watershed. Specifically, I: 1) characterized the breeding structure of the cohort,

with regard to the number of spawners that produced it, the mating system of the spawners, and the

genetic characteristics of juveniles versus parents, 2) estimated the spatial distances potentially dispersed

by juvenile fish, and 3) estimated mean lifetime dispersal distance. I hypothesized that dynamic stream

conditions during spring spawning would result in low reproductive success for spawners, and thus that

most juveniles would come from relatively few families (e.g., Hudy et al. 2010). I further hypothesized

that real-time dispersal by juveniles would be spatially restricted, as reflected in spatial clustering of

family members (from PR) and a lack of long-distance cross-assignments (in ATs). However, I

hypothesized that gene flow synthesized over many generations would be spatially extensive (e.g.,

throughout the watershed), due to the cumulative effect of many short "stepping stone" dispersal events.

METHODS

Study species and area

The Roanoke logperch (*Percina rex*) is a large-bodied (to 165 mm total length) member of the

darter subfamily (Percidae: Etheostomatinae). The species is iteroparous, with overlapping generations,

age-at-maturity of 2-3 years, and lifespan of up to 6.5 years (Burkhead 1983). The mating system is

unknown, but like many other darters (Page 1983), could involve both male and female polygamy. In the

Roanoke River population, spawning occurs in April-May over swift gravel runs; eggs are buried with no

subsequent parental care (Burkhead 1983). Young occupy shallow pool margins and backwaters until

September-October, when they begin to shift into the riffle-run habitats occupied by adults (Rosenberger

and Angermeier 2003; Roberts and Angermeier 2010). These discrete riffle-run patches can be considered

local breeding habitats, and often are separated from each other by long (> 1 km) sections of unsuitable

pool habitat; movement among riffle-runs therefore constitutes dispersal under the definition provided

above.

The geographic range of *P. rex* includes approximately 102 river km of the upper Roanoke River

watershed, including the mainstem Roanoke River and its North and South forks (Figure 2.1). Upstream

distributional limits of this population are determined by stream size, whereas the downstream limit is

marked by Niagara hydroelectric dam, a large barrier to fish movement. No other known year-round

movement barriers are present in the watershed, but various low-water bridge crossings and culverts

probably are impassable during low flows. Upstream portions of the watershed drain predominantly

forest and farmland, whereas the downstream one third of the watershed drains heavily urbanized areas of

Salem, Roanoke, and Vinton, VA. As is typical of a Valley and Ridge system, the watershed topology is trellised, with small steep tributaries emptying into large main channels. At the downstream end of the study area, stream temperature ranges from 0 to 28 °C annually, and mean daily discharge ranges from 3.6 to 22.7 m^3 s^{-1} annually (data from U.S. Geological Survey, http://waterdata.usgs.gov/va/nwis/uv/?site_no=02055000).

Sample collection and processing

I collected *P. rex* from 15 sites throughout the watershed, between July and September of 2005 and 2006 (Table 2.1; Figure 2.1). Each site was 100-300 m long and ranged from 12 m (Site NF2) to 34 m (Site RR1) wide. I obtained geographic coordinates at the midpoint of each site using a handheld GPS receiver. Sites were separated by 1.2 to 54.7 km of stream (mean = 20.8 km). Fish were captured with a Smith-Root direct-current backpack electrofisher and seines or dipnets. The total length (mm) of each fish was recorded, a small section of tissue was excised from the caudal fin, and then fish were returned alive to the stream. Tissue samples were dried individually in paper envelopes and stored at -20 °C until DNA extraction.

I sorted fish into age classes based on length-frequency histograms. I assumed that all Age-2+ fish potentially were mature (although some females do not mature until Age 3; Jenkins and Burkhead 1994) and therefore that fish Age 2 or older in 2005 samples or Age 3 or older in 2006 samples were potential parents of the 2005 cohort. Sample sizes varied among sites and years due to the general rarity of the species and likely temporal variation in recruitment. The sample of candidate parents likely comprised multiple age-classes but was not sorted into cohorts because the precise ages of Age-3+ fish were difficult to determine based on length.

DNA isolation and genotyping

I extracted template DNA from whole tissue samples using a PureGene DNA Extraction Core Kit A (Gentra Systems, Minneapolis, Minnesota, USA). I genotyped samples at eleven microsatellite DNA loci (*Prex33*, *Prex36*, *Prex37*, *Prex38*, *Prex41*, *Prex42*, *Prex43*, *Prex44*, *Prex45*, *Prex46*, and *Prex47*) developed for *P. rex* by Dutton et al. (2008) using PCR conditions reported therein. Forward primers for

46

these loci were labeled with NED, VIC, PET or FAM fluorescent dye (Applied Biosystems, Inc., Foster City, California, USA), and PCR was conducted in a MyCycler Thermal Cycler (BioRad, Hercules, California, USA). Amplification products were separated using an ABI 3130 automated sequencer and sized using GENEMAPPER 3.5 and a LIZ500HD size standard (Applied Biosystems, Inc., Foster City, California, USA).

Genetic diversity and differentiation of age groups

To ensure that no individuals had been sampled twice, I searched for matching multilocus genotypes using GenAlEx 6.2 (Peakall and Smouse 2006); samples with genotypes identical at > 10 loci were considered to be from the same individual. I tested for Hardy-Weinberg equilibrium for the 2005 cohort (subsequently "juveniles") and candidate parents of this cohort (subsequently "adults") both separately and combined, as well as for linkage equilibrium for all samples combined, using ARLEQUIN 3.11 (Excoffier et al. 2005). Hardy-Weinberg tests employed 10^5 recorded Markov-Chain Monte Carlo (MCMC) chains following a burn-in of 10^3 chains, whereas linkage tests employed 10^5 randomizations. Results were evaluated using a sequential Bonferroni adjustment for a global $\alpha = 0.05$. I also tested for genotyping errors (i.e., null alleles, large allele dropout, and stutter) within the juvenile sample using MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004).

I asked whether levels of genetic diversity were similar between juveniles and adults by comparing unbiased gene diversity (H_E), observed heterozygosity (H_O), the inbreeding coefficient (F_{IS}), and allele richness estimated using FSTAT 2.9.3.2 (Goudet 2002). The magnitude of genetic differentiation between juveniles and adults was estimated using Weir and Cockerham's (1984) estimator of F_{ST}, calculated in SPAGeDi 1.3 (Hardy and Vekemans 2002); the null hypothesis that $F_{ST} = 0$ was tested via 10^4 permutations of individuals between cohorts. To further examine potential sources of genetic variation within the sample of juveniles, I used an analysis of molecular variance (AMOVA) in ARLEQUIN to decompose total genetic variance into the following sources: capture year (2005 versus 2006), sampling sites within capture years, and residual variation among individuals within sites. The

magnitude of variance due to each source was tested for equality with zero using 10^4 random permutations of objects among groups at the stratum being tested.

I estimated the genetically effective size of the upper Roanoke River *P. rex* population using various approaches. First, I estimated the number of breeders that produced the 2005 cohort (N_b) using Waples' (2006) bias-corrected linkage disequilibrium (LD) method, as implemented in LDNe (Waples and Do 2008). I excluded rare alleles occurring at frequency < 0.05 and estimated the mean and 95% confidence interval of N_b by jackknifing over loci. Given that the PR suggested that promiscuity was common in *P. rex* (see Results), I assumed random mating rather than monogamy in the LDNe model. Second, I estimated N_b using Wang's (2009) method based on the proportion of reconstructed full- and half-siblings in the sample, as implemented in COLONY 2.0 (Jones and Wang 2010). Third, I estimated average effective population size (N_e) across the generations immediately preceding the 2005 cohort using the approximate Bayesian computation (ABC) method implemented in ONeSAMP (Tallmon et al. 2008). I ran five replicate ABC models, each consisting of 5×10^4 simulated populations and assuming a uniform prior distribution on N_e ranging from 2 to 10^4 individuals. I then weighted each replicate's estimated mean N_e by the width of its 95% credible interval and computed a grand harmonic mean N_e across replicates.

Pedigree reconstruction

I used the group-likelihood pedigree approach of Wang and Santure (2009), as implemented in COLONY 2.0, to reconstruct families of *P. rex*. In the approach, sampled offspring are clustered into hypothetical maternal and paternal families, and then candidate parents are assigned (or not assigned) to these families. Models assume that a random sample has been collected from a population at Hardy-Weinberg and linkage equilibrium. Uncertainty associated with family assignments is estimated by integration across intermediate-likelihood configurations and can be lessened by averaging across replicate model runs (Wang and Santure 2009). The approach can accommodate monogamous, one-sex-polygamous, and promiscuous mating systems. It also can incorporate genotyping errors due to null alleles (Class I) or random processes such as mutation, stutter, or miscalling (Class II; Wang 2004).

The accuracy of PR presumably depends strongly on sample size, the number and variability of loci, and the choice of a mating model appropriate to the species (Wang 2004), although the effects of the latter are poorly investigated. As an indication of the statistical power of the set of loci for inferring pedigree information, I estimated the probabilities that two randomly selected non-related individuals (PI_{NON}) or siblings (PI_{SIBS}) would have identical genotypes using GenAlEx. Further, to estimate my potential accuracy for reconstructing the unknown family structure of *P. rex* in 2005, I assessed the accuracy of family assignments made on simulated datasets with known pedigree. To do this, I randomly selected 14 male and 14 female sampled adult *P. rex* genotypes to use as simulated parents, and then created simulated offspring by "pairing" these adults in various mating strategies. Each simulated offspring's genotype was constructed by randomly selecting one allele from each parent independently for each of the 11 loci.

Four simulated datasets with known pedigree were created: (1) *Dataset 1 - Monogamy*: Each of 14 males was paired three times with one other female, producing 14 full-sib families containing 3 offspring each. (2) *Dataset 2 - Polygyny*: Each of 7 males was paired twice with each of 2 different females, producing 14 full-sib families containing 2 offspring each, nested within 7 half-sib families containing 4 offspring each. (3) *Dataset 3 - Promiscuity*: Each of 14 males was paired twice with each of two different females and vice versa, producing 28 full-sib families containing 2 offspring each, nested within 7 "clusters" containing 8 offspring each. Each cluster contained 4 full-, 16 half-, and 8 non-sib pairs of offspring. (4) *Dataset 4 - All offspring unrelated*: Each male was mated once with one other female, producing 14 full-sib families containing 1 offspring each.

I analyzed simulated datasets in COLONY 2.0 using the medium run length and full-likelihood estimation method with medium precision, combining likelihoods over three replicate model runs. Each dataset was analyzed using twelve different modeling strategies, consisting of all factorial combinations of two assumed mating systems (monogamy versus promiscuity), two levels of availability of parental genotype data (included versus excluded), and two assumed rates of Class I and II genotyping errors (0 versus 0.05 for all loci). In analyzing model results, I considered two individuals to be "matched" if the

probability of the relationship (sibling-sibling or parent-offspring) was $\geq 95\%$. I calculated model sensitivity and specificity based on the proportions of correctly matched pairs and correctly unmatched non-pairs, respectively.

I reconstructed the unknown 2005 family structure of *P. rex* in COLONY 2.0 using the modeling intensity described above. Based on results of simulation studies (see Results), I conservatively allowed for polygamy of both sexes but no genotyping errors. Four offspring were removed from this analysis because of incomplete genotypes. Thus, the data consisted of 90 offspring, 17 candidate fathers, and 19 candidate mothers. Once family groups were constructed, the known capture locations and dates of *P. rex* individuals were used to estimate the spatial dispersion of family members over time. Because juveniles were captured at few sites in 2005 (Table 2.1), maximum detectable displacement distances varied from 18 km in 2005 to 55 km in 2006 (Figure 2.2). To determine whether dispersion of juveniles increased over time, I used a nonparametric resampling-based ANOVA in RESAMPLING PROCEDURES 1.0 (10^5 permutations; Howell 2000) to test whether mean displacement distance of siblings differed between a) siblings both captured in 2005, b) sib pairs in which one member was captured in 2005 and the other in 2006, and c) siblings both captured in 2006. For the sake of comparability, only captures made at the six sites where juveniles were caught in both years (Table 2.1) were used for the ANOVA. Furthermore, to test whether the dispersion of family members was more spatially restricted than predicted by chance, I compared the distribution of spatial distances separating capture locations of inferred siblings to the distribution of distances separating inferred non-siblings. My rationale was that if dispersal is restricted, juveniles will be more likely related to nearby than distant individuals. I compared distributions using a randomization-based *t*-test for difference in central tendency in RESAMPLING PROCEDURES and a Fisher's exact contingency-table test for difference in shape in R 2.10.0 (The R Foundation for Statistical Computing 2009).

Assignment tests

I used individual assignment tests (ATs) to assess whether juveniles were more likely to have originated from a site other than the one where they were captured, and whether rates of these "cross-

assignments" between sites were negatively related to the distance separating sites. I presumed that higher rates of cross assignment indicated more frequent dispersal between sites. Castric and Bernatchez (2004) found that spatial analyses of such cross assignments were more powerful than traditional IBD tests for detecting spatially restricted dispersal, possibly because ATs make use of full multilocus genotypes, whereas summary statistics (e.g., F_{ST}, $\hat{a}$) are calculated on a locus-by-locus basis and therefore contain less information. I estimated each juvenile's likelihood of origination from each of the 13 sites at which two or more juveniles were collected using Rannala and Mountain's (1997) Bayesian assignment method implemented in GENECLASS 2.0 (Piry et al. 2004). The site with the highest likelihood was presumed to be that individual's site of origination, and the cross-assignment rate between site pairs was calculated as the number of individuals cross assigned between the pair of sites divided by the total number of individuals captured at the two sites. I tested for an upstream versus downstream bias in cross-assignment rates using a Fisher's exact contingency-table test in R. I then assessed the relationship between cross-assignment rate and the spatial distance separating sites over all distance classes (i.e., 1-55 km), as well as at cumulatively increasing 5-km increments (i.e., 1-5 km, 1-10 km, etc.), by estimating the mean and bootstrapped 95% confidence limits of Pearson's correlation coefficient (r) based on 10^4 re-samples in RESAMPLING PROCEDURES. As an index of my statistical power to detect migrants, I estimated the likelihood ratios of observing each genotype in its site of capture versus each other site in turn (i.e., D_{LR} values) following Paetkau et al. (1997).

Lifetime dispersal distance

Wright (1943) showed that in a species evolving under isolation-by-distance (IBD), the rate of decrease of correlation of gene frequencies over space is inversely related to the "neighborhood size" ($D\sigma^2$), where D is the density of effective breeders per unit distance and σ is the average lifetime dispersal distance (i.e., between natal and breeding locations). Rousset (1997; 2000) further showed that under one-dimensional IBD, there is a positive linear relationship at equilibrium between the genetic differentiation and spatial distance separating pairs of demes or individuals, and that the regression slope (β) of this relationship is a linear function of neighborhood size, as follows:

51

Equation 1: $\beta = (4D\sigma^2)^{-1}$

I estimated β empirically using IBD data from the 2005 cohort. If individuals are distributed among discrete demes or are widely separated, a group-based measure of differentiation is more appropriate for analyzing IBD [i.e., F_{ST} / (1- F_{ST}); Rousset 1997], whereas if individuals are distributed more uniformly in space, an individual-based measure of differentiation is more appropriate [i.e., $\hat{a}$; Rousset (2000)]. Lacking knowledge of the more appropriate formulation for *P. rex*, I performed IBD regressions using both measures. I estimated means and 95% confidence intervals of β by jackknifing over loci in SPAGeDi. Because study-area extent can strongly affect the estimate of β (Leblois et al. 2003), I analyzed IBD over the full range of pairwise distances (1 to 55 km) as well as over several reduced distance ranges (i.e., 1-25 km, 1-40 km, 5-55 km, and 10-55 km). Zero-distance (i.e., same-site) comparisons were not included in any of the above sets, and only sites at which two or more juveniles were collected (*n* = 13) were used to estimate F_{ST}. To assess whether loci with high diversity biased β downward (e.g., Leblois et al. 2003), I estimated the correlation among loci of β with H_E using R.

Extraction of σ from estimates of neighborhood size required an estimate of *D*. I developed three estimates of *D* for *P. rex* by dividing each of my three estimates of the effective population size (see above) by the total extent of the study area, 66.9 km. This approach assumed panmixia and constant density of spawners throughout the study area.

RESULTS

Genetic diversity and differentiation

In total, I collected DNA samples from 94 juveniles and 36 candidate parents (17 M:19 F) of the 2005 cohort (Table 2.1). No individuals exhibited matching genotypes at more than three loci, indicating that no individuals had been sampled twice. Considering all samples, no loci exhibited significant evidence for linkage disequilibrium, nor did any loci exhibit significant evidence for Hardy-Weinberg disequilibrium within offspring, parents, or both groups combined. Furthermore, no loci exhibited evidence for genotyping errors due to stuttering, large allele dropout, or null alleles.

Estimates of genetic diversity were identical between juveniles and parents: H_E, H_O, F_{IS}, and A_{36} (allele richness standardized to 36 individuals) were 0.83, 0.80, 0.04, and 10.0, respectively, for both groups. The estimated magnitude of genetic differentiation between juveniles and parents was low ($F_{ST} = 0.001$) and not statistically distinguishable from zero ($P = 0.27$). Within the sample of juveniles, the majority of genetic variation (99.2%; $P < 0.0001$) was among individuals within sample sites, whereas variation between capture years and among sampling sites within years was trivial and statistically non-significant (0.5 and 0.3%, respectively; both $P > 0.05$).

Estimates of effective population size varied widely among methods. Mean (95% confidence limits) N_b values from the LD and COLONY methods were 1218 (339-infinity) and 105 (78-148) individuals, respectively. Mean N_e from the ABC method was intermediate to the other two methods, with a weighted posterior mean (outer 95% credible limits) of 318 (110-2205) individuals.

Statistical power of loci and models for pedigree reconstruction

Probabilities of identity indicated that the 11 microsatellite loci were sufficiently variable to provide high resolution of pedigrees. Within the sample of juveniles, estimated probabilities of identity for non-related individuals and siblings were 7.8×10^{-16} and 8.7×10^{-6}, respectively. Even with data from only 10 loci, PI_{NON} and PI_{SIB} were at least 5.3×10^{-13} and 2.9×10^{-5}, respectively.

I evaluated the accuracy of COLONY 2.0's PR approach using simulated datasets with known family pedigrees (see Appendix C). Overall, models were conservative, in that they sometimes failed to match related family members but never incorrectly matched unrelated pairs of individuals. Analyses also indicated the importance of assuming the correct mating system when performing PR. Model specificity (i.e., the percentage of correctly unmatched non-family-pairs) was 100% in all 32 models. However, model sensitivity (i.e., the percentage of correctly matched family pairs) varied across models from 16% to 100%. Sensitivity depended on the type of pair being assessed (i.e., sibship, parentage), as well as on the actual and assumed mating system, whether or not data from parents were included, and the assumed rate of genotyping errors. If the true mating system was monogamous, model accuracy was highest if monogamy was assumed. However, if any type of polygamy occurred, a monogamy model was

less accurate than a promiscuity model. When a monogamy model was applied to a polygamous mating system, half-sibs were incorrectly assigned either as full-sibs or unrelated pairs. Parentage results followed these same general patterns, and sibship assignments were more accurate when parental genotypes were available to models. Allowing for genotyping error rates had little substantive effect on model results, and models with error included did not consistently exhibit more or less accuracy than models without error. Given these findings, I took a conservative approach to PR for *P. rex*, allowing for promiscuity but no genotyping errors.

Pedigree reconstruction for the 2005 cohort

Based on PR, the majority of *P. rex* sampled in 2005 and 2006 were unrelated. The sample of 90 juveniles consisted of zero full-sib pairs, 61 half-sib pairs, and 3944 non-sib pairs. Of the 36 candidate parents, one candidate father was assigned to one offspring and no candidate mothers were assigned to offspring. Juveniles had an average of 1.4 sampled half-siblings (range 0 to 4) and reconstructed male and female parents each had an average of 2.4 mates (range 1 to 5). Thus, I found evidence for promiscuity of both sexes and the existence of many small families.

The two members of the single deduced father-offspring pair both were captured at Site 4 (the juvenile in 2005, the father in 2006). In contrast, the majority of deduced sibling pairs (90%) were dispersed across multiple sites (Figure 2.2). Of the 61 half-sib pairs, only 9 involved juveniles that both were captured in 2005. These captures, which represented spatial dispersion within the first 2-5 months of life, were separated by an average of 5.5 km (range 0.0 to 13.2 km). In the remaining 52 half-sib pairs, one or both members of the pair were captured in 2006. These captures, which represented spatial dispersion within the first 13-16 months of life, were separated by an average of 15.5 km (range 0.0 to 54.7 km). Controlling for differences in sampling intensity between years, there was no significant difference over time in the mean displacement distance separating siblings ($F = 0.16$; $P = 0.85$). Siblings more frequently were separated by short than long displacement distances (Figure 2.2). However, the mean displacement distance separating siblings (14.0 km) was not significantly different from the mean distance separating non-siblings (15.5 km; $t = 0.75$; $P = 0.46$), and the displacement distributions for

siblings and non-siblings were not significantly different in shape (Fisher's exact $P = 0.16$). Thus, I could not reject the hypothesis that *P. rex* siblings were distributed randomly throughout the study area.

Assignment tests

The majority of the 92 juveniles (93%) analyzed in ATs were cross-assigned to a site other than the one in which they were sampled, indicating weak overall genetic structure among sites. Consistent with this lack of structure, the mean D_{LR} among sites was 1.8 (SE = 0.12), which indicated that my statistical power to distinguish migrants from residents was low (i.e., < 0.5; Paetkau et al. 2004). Therefore, I cautiously interpreted cross-assignments as indices of dispersal patterns rather than dispersal events *per se* (Castric and Bernatchez 2004). Among cross-assigned individuals, the assigned site of origin was more often downstream than upstream of the site of capture (47 vs. 39 cases, respectively), but this bias was not statistically significant (Fisher's exact $P = 0.22$). I therefore ignored assignment direction for subsequent analysis. As expected under spatially restricted dispersal, when all distance classes were considered, the relationship between cross-assignment rate and spatial distance was negative ($r = -0.19$) and the 95% confidence limits of r fell below zero (Figure 2.3). However, cross-assignment rates did not decrease with increasing distance until sites were separated by ≥ 20 km, and the confidence limits of r overlapped zero until sites were separated by ≥ 50 km.

Lifetime dispersal distance

Genetic differentiation of juveniles was weak between sites (mean $F_{ST} = 0.004$) and between individuals located in different sites (mean $\hat{a} = 0.038$), but in both cases was positively related to spatial distance (Table 2.2). The estimated slope of IBD was similar regardless of whether group- ($\beta = 0.00025$) or individual-centered ($\beta = 0.00029$) differentiation statistics were employed. Estimates of β over reduced distance ranges were of the same order of magnitude (range = 0.00024 to 0.00056) as those based on the full range, with the exception of one that was negative (i.e., F_{ST} over the 1-25 km range). Furthermore, although the genetic diversity (H_E) of individual loci varied from 0.68 to 0.92, there was no significant correlation among loci of diversity with β, whether β was estimated from $\hat{a}$ ($r = 0.34$, $P = 0.31$) or F_{ST} ($r = 0.41$, $P = 0.22$). I therefore rejected the hypothesis that high-diversity loci biased mean IBD

slopes downward. However, 95% confidence intervals of β were wide and overlapped with zero in all

cases. Negative β results in a nonsense estimate of neighborhood size, so the lower bounds of β could not

be used to estimate the upper bounds of σ; however, the upper bounds of β always were positive and

therefore could be used to estimate the lower bounds of σ.

Uncertainty about the most appropriate estimator of differentiation (F_{ST} versus $\hat{a}$) had little effect

on demographic estimates. However, variation in effective population size (N_e or N_b) across estimators

(outer 95% confidence limits from 78 to 1218 individuals) and β across loci (outer confidence limits from

zero to 0.00181) resulted in wide uncertainty around estimates of effective breeder density (*D*) and

lifetime dispersal distance (σ; Figure 2.4). Given the confidence limits of effective population size, the

derived 95% confidence limits of *D* were estimated to be 1.2-18.2 breeders km^{-1}. Based on an average

estimate of β (= 0.00025, based on the F_{ST} regression over all distance classes), the derived 95%

confidence limits of σ were estimated to be 7.4-28.9 km. Uncertainty about the upper limit of β had little

influence on σ, whereas β values near the lower limit resulted in large values of σ (Figure 2.4).

DISCUSSION

Breeding structure and effective size of the 2005 cohort

My genetic investigation of the upper Roanoke River population of *P. rex* provided insights into

reproductive biology that would have been difficult to obtain using other means. One noteworthy finding

based on pedigree reconstruction (PR) was that most juveniles that I captured were not closely related, but

instead derived from many small families. The widespread success of spawners in 2005 prevented

genetic drift from producing appreciable differences in allele frequencies between the parental and

offspring cohorts; parent and offspring samples were genetically indistinguishable and exhibited high,

identical levels of genetic diversity. Spawners of both sexes likely enhanced their reproductive success

by adopting a promiscuous mating strategy. PR indicated that both males and females mated with an

average of 2.4 partners in 2005. Promiscuity has not previously been documented for *P. rex*, but is not

surprising given that groups of multiple males and females aggregate on the spawning grounds (Jenkins

and Burkhead 1994) and given that other egg-burying darters, including other logperch species, mate with multiple partners (Winn 1958; Page 1983).

Variation among my estimates of N_b was unexpectedly large, but may reflect idiosyncrasies of the different methods I used. Depending on the method, estimated N_b numbered in the hundreds to thousands of individuals. Under the sampling scheme of this study (approximately 100 samples and 10 loci), the linkage disequilibrium (LD) method is upwardly biased and the sibship reconstruction (COLONY) method downwardly biased when true N_e is $\geq$ 100 individuals, and the severity of these biases increases with further increases of N_e (Wang 2009; Waples and Do 2010). Under these conditions, COLONY also exhibits overly narrow confidence limits. The Approximate Bayesian Computation (ABC) method has not been thoroughly evaluated for bias or precision, but in theory should be the least biased and most precise estimator because it takes advantage of more types of data (Luikart et al. 2010). However, a disadvantage of the ABC method is that the N_e it estimates is not directly comparable to the N_b estimated by the two other methods. In my case, the ABC method estimated the number of individuals necessary to maintain for 2-12 generations a set of genetic characteristics similar to my sample of the 2005 cohort (Tallmon et al. 2008), whereas the LD and COLONY methods estimated the effective number of breeders that produced the 2005 cohort. If N_e varies widely from generation to generation, then these quantities may differ substantially. Furthermore, all three methods assume that generations do not overlap, which complicates their interpretation for an age-structured species like *P. rex* (Luikart et al. 2010; Waples and Do 2010). Lacking an ideal estimator of N_b for *P. rex*, I used multiple methods to develop a reasonable range of estimates and incorporated uncertainty across methods into my confidence limits of N_b and σ.

Independent, demographic estimates of effective population size (*sensu* Caballero 1994) are lacking for this population, but genetic estimates are consistent with what we know about the abundance of potential spawners in 2005. During concurrent sampling of Roanoke River riffles in 2005 for another study, I observed an average of 1.7 Age-2+ *P. rex* individuals per riffle (Roberts and Angermeier 2010). However, my estimated sampling efficiency is only ~10% (Roberts and Angermeier, unpublished data), so the true abundance of spawning-age fish probably was closer to 17 fish per riffle. Burkhead (1983)

counted a total of 472 riffles within the known range of *P. rex* in the watershed. If I assume constant

abundance of spawners across riffles, then I roughly estimate that there were 8,024 spawning-age fish in

the population in 2005. This calculation may overestimate abundance, if *P. rex* density is lower in the

North and South forks than in the mainstem Roanoke River (Burkhead 1983). Nevertheless, the range of

N_b:N ratios (0.01-0.48) that I derive from my genetic estimates of N_b and direct counts of *P. rex* in 2005 is

within the range of N_e:N ratios previously observed across taxa like fishes that exhibit type-III

survivorship curves (Frankham 1995; Turner et al. 2006; Palstra and Ruzzante 2008).

Juvenile dispersal

My previous work indicated that *P. rex* exhibits high gene flow across watersheds (see Chapter

1), but I did not know whether extensive gene flow was accomplished via single- or multiple-generation

dispersal, or whether patterns of dispersal vary over ontogeny. In this study, I used assignment test (AT)

and PR methods to directly measure juvenile dispersal by identifying migrant individuals. ATs are most

powerful when allele frequencies differ substantially among potential source populations, but have low

power to distinguish migrants from residents when F_{ST} is < 0.02 (Hauser et al. 2006; Hall et al. 2009)

and/or D_{LR} is < 3 (Paetkau et al. 2004) between populations. In the case of *P. rex* juveniles, sample sites

exhibited weak allele-frequency divergence (mean F_{ST} = 0.004 and D_{LR} = 1.8), which likely limited the

ability of GENECLASS2 to build discriminatory likelihood functions. Moreover, low sample sizes

limited the precision with which each site's reference allele frequencies could be characterized.

Exclusion tests can be used to assess whether assignments exceed a threshold probability necessary to

conclude migrant ancestry (Manel et al. 2005; Hall et al. 2009), but I had no objective means for setting

such a threshold. As a result, I could not confidently discern true current-generation dispersal from recent

mixing in past generations that produced "assignment errors" (Waser and Strobeck 1998). Thus, rather

than interpret AT cross-assignments as dispersal events *per se*, I interpreted rates of cross-assignment as

relative indices of dispersal between sites (Castric and Bernatchez 2004).

Caveats notwithstanding, the AT suggested frequent dispersal of juveniles between sites, with

proportions of cross-assigned individuals ranging from 0 to 36% for individual site-pairs. Apparent

dispersal was particularly frequent between sites separated by ≤ 15 km; over this extent, individuals were as likely to be cross-assigned to non-adjacent as adjacent sites. Over greater distances (≥ 20 km), cross assignments declined with increasing distance, but occurred even at the largest distance class (50-55 km). This finding contrasts with those of Castric and Bernatchez (1994), who observed that for riverine brook trout *Salvelinus fontinalis* and Atlantic salmon *Salmo salar*, the relationship between cross assignment rate and distance was stronger at smaller than larger spatial extents. More generally, both theoretical (Rousset 2000; Leblois et al. 2003) and empirical studies (Hänfling and Weetman 2006; Whiteley et al. 2006) often observe a negative relationship between the geographic extent of a study and the strength of IBD, presumably because the spatial effects of dispersal on genetics are overwhelmed at large spatial extents by the effects of drift, mutation, and sampling error. On the contrary, I observed no weakening of IBD with geographic extent in *P. rex*, whether IBD was gauged using AT cross-assignment rates or the genetic distance measures F_{ST} and $â$ (which were similarly related to distance over all extents considered). I attribute this contrast to extensive dispersal by *P. rex*, which apparently exerts significant influence (relative to drift) throughout the spatial extent covered by this study. Observed extensive dispersal by *P. rex* contrasts sharply with the conventional wisdom that stream fish seldom move beyond reach boundaries (e.g., Gerking 1953) and argues for a watershed-grained focus for monitoring and management of the species (Fausch et al. 2002).

Like the AT, PR indicated extensive dispersal of *P. rex* throughout the Roanoke River watershed. Based on capture locations of juveniles, deduced Age-0 and Age-1 half-siblings were separated by up to 13 and 55 km, respectively. Though these figures are not directly comparable because of the narrower sampling extent in 2005 than 2006, they indicate spatially broad and temporally persistent dispersion of juvenile *P. rex*. Like other stream fishes (e.g., Hall 1972; Northcote 1978; Humston et al. 2009; Morissey and Ferguson 2010), many darters undergo a life-cycle comprising primarily upstream spawning migration by adults which counteracts subsequent downstream movement of larvae and juveniles (Turner 2001; Slack et al. 2004; Roberts and Angermeier 2007). Ontogenetic variation in the longitudinal positions preferred by younger versus older fish reflects tradeoffs between the productivity, stability,

volume, and safety from predation offered by upstream versus downstream habitats (Schlosser 1987). Young-of-year fish, particularly when < 10 mm TL, also are susceptible to passive downstream dispersal during floods (Harvey 1987). However, there were no high-flow events in the Roanoke River between 15 April and 30 June 2005, when the 2005 cohort would have been most vulnerable to displacement (Jenkins and Burkhead 1994). Furthermore, there was no significant directional bias to AT cross-assignment rates, and the two members of the only inferred parent-offspring pair were captured at the same site. In most cases, displacement distance between sibling pairs could not be assigned a dispersal direction because I did not know which site, if either, was the natal habitat. However, in the few cases where family members were captured across multiple years, the temporal sequence of capture locations did not indicate a directional tendency.

An alternative explanation for the observed extensive dispersion of siblings is that parents spawned at multiple sites over the course of the spawning season. It seems reasonable to assume that full siblings originate from the same spawning event and location, and therefore that spatial displacement between full sibs represents natal dispersal (e.g., Danancher et al. 2008; Hudy et al. 2010). However, I did not capture any full siblings. In contrast, there is no reason to assume that the half siblings I captured originated from the same spawning location, or even from the same riffle. Large *Percina* species are active swimmers that roam throughout riffles to feed (Greenberg 1991; Roberts et al. 2008), exhibit little spawning-site fidelity to particular portions of riffles (Winn 1958), and migrate extensively up- and downstream among riffles over the course of a spawning season (Winn 1958). Therefore, *P. rex* half siblings may have been spawned at multiple sites as their shared parent mated with various partners.

For logistical reasons, past CMR studies of darter movement have focused on adults and large juveniles (Roberts and Angermeier 2007). Such studies generally have found spatially limited movement, though most studies were not extensive enough to detect dispersal events (reviewed in Schwalb et al. 2011). Another explanation for the limited movement observed is that fish performed much of their dispersal during the unstudied small-juvenile phase. This appears to be the case for smallmouth bass *Micropterus dolomieu*, which have been shown to exhibit high spawning-site fidelity as adults (Gerking

1953; VanArnum et al. 2004) but extensive natal dispersal as young-of-year (Humston et al. 2009). In contrast, resident brown trout *Salmo trutta* were shown to exhibit limited dispersal as young-of-year, but gradually increased dispersion with age (Vera et al. 2010), whereas studies of brook trout have observed both patterns (Hudy et al. 2010; Kanno et al. 2011; Morissey and Ferguson 2010). For fixed ontogenetic differences in dispersal to persist, they must confer a net fitness benefit to dispersing individuals. Potential benefits of dispersal include avoidance of inbreeding and competition with close relatives and access to new habitats where resources are more common, whereas potential costs include increased energetic expenditures and vulnerability to predation (Ronce 2007). I suspect that the importance of these benefits and costs varies widely across space and time in streams, and thus may vary as much within as between fish taxa.

Clearly, a more thorough understanding of the ontogeny of dispersal, and its variability across fish populations, is needed. Regardless of whether spatial displacement of *P. rex* siblings was due to breeding dispersal, natal dispersal, or both, it resulted in watershed-scale redistribution of progeny. This spreading of reproductive output across locations may allow an individual to hedge bets against environmental variability in dynamic stream environments (Winemiller and Rose 1992).

Lifetime dispersal distance

The reproductive success of mobile versus resident individuals is scarcely investigated for stream fishes (but see Hendry et al. 2004), yet work on other taxa suggests that immigrants often fail to establish or successfully reproduce in their new habitat (Greenwood 1980; Nosil et al. 2005; Hall et al. 2009). Furthermore, although natal homing is well known for salmonines (Northcote 1978) and some centrarchids (Ridgway and Shuter 1996), its prevalence among other stream fishes, including darters, is unknown. Direct dispersal-estimation methods like ATs and PRs can measure the spatial displacement of a group of individuals, but cannot inform whether displaced individuals ultimately spawn, and if so, whether they remain in their new habitat or return to their natal habitat to do so. If extensively-dispersed *P. rex* juveniles ultimately fail to spawn, or if they re-aggregate in natal habitats prior to spawning, then my direct estimates of dispersal would overestimate the geographic extent of effective connectivity

61

(Greenwood 1980; Lowe and Allendorf 2010). Therefore, for comparison to direct estimates, I used indirect equilibrium models to estimate the mean lifetime dispersal distance between natal and spawning locations.

Although there was considerable uncertainty around parameter values, my estimates of mean lifetime dispersal distance (7-29 km) bracketed both the mean distance separating inferred siblings (14 km) and the spatial domain of high AT cross assignment ($\leq$ 15 km). This concordance across multiple methods suggests that: a) dispersal patterns of the 2005 cohort were broadly representative of those of an average cohort, synthesized over the long term, b) the species performs most of its lifetime dispersal during the juvenile phase, c) extensively-dispersed individuals do, on average, successfully reproduce in new habitats, and d) *P. rex* should exhibit high demographic and genetic connectivity throughout the Roanoke River watershed. Interestingly, this dispersal distance also is similar to the spatial extent (~10-20 km) over which Hitt and Angermeier (2008) detected influences of riverine immigrants on local stream fish assemblages (which did not include *P. rex*), suggesting that fish movement over such scales may be relatively common.

There is a surprising lack of corresponding estimates of lifetime dispersal distance for other stream fishes, given the apparent applicability of the IBD-slope method to stream environments. In the only other application of which I am aware, Koizumi et al. (2006) estimated the σ of a migratory Dolly Varden char population to be only 1.4-2.5 km. This finding supports the notion that, although they move great distances over a lifetime, migratory salmonines generally exhibit strong natal homing to spawn (Northcote 1978). In contrast, the large lifetime dispersal distance exhibited by *P. rex* highlights the potential diversity of spatial-ecological strategies expressed among stream fishes and the need of further research on dispersal for non-salmonine taxa.

Utility of genetic methods for estimating dispersal of stream fishes

Stream fish dispersal is intrinsically difficult to study; it requires knowledge of the locations of cryptic organisms at various times of their life cycle within an environment that is difficult for humans to access and sample. I feel that the analysis of molecular genetic data provides useful solutions to these

challenges, both supplementing and complementing traditional CMR approaches (Lowe and Allendorf 2010). For example, genotypes can be used as individual-specific "tags", and spatial data from individuals captured multiple times can be analyzed in a traditional CMR framework. In this study, I used methods that require only one capture per individual and seek to draw inferences about dispersal from the nature of the genetic variation itself.

Of the methods I used, PR probably offers the most promise for increasing our insights into stream fish dispersal, as well as other aspects of population biology (Morrissey and Ferguson 2011). Analyses of simulated data indicated that COLONY 2.0's algorithm was accurate at resolving complex family relationships involving polygamy of both sexes. The method also was conservative, in that it sometimes failed to match true family members but never falsely matched unrelated individuals. For most applications, this type of error is preferable to the opposite type; false matching of disparate individuals likely inflates estimates of dispersal, whereas failure to recognize some family pairs injects no systematic biases into dispersal estimates. In any event, accurate sibship reconstruction is important, and I found that it was dramatically improved by including samples from parents and by assuming an appropriate mating system. Capturing both the parents and offspring of families may be difficult in many field studies, but is an important goal. Likewise, the mating systems of many species are yet unknown, but my simulations suggest that unless there is evidence for strict monogamy, the safest analytical strategy is to allow for promiscuity. Otherwise, half-siblings will be incorrectly assigned as full- or non-sibs and demographic parameters will be inaccurately estimated. Findings from studies that apply a monogamy model to a polygamous species should therefore be viewed with caution. Finally, I advise that in future studies that use PR to measure juvenile dispersal, the initial sampling of individuals should take place as soon as possible after hatching, so that the natal site can be pinpointed and the directionality and distance of subsequent natal dispersal events can be assigned.

ATs also show promise for measuring dispersal, albeit under a narrower range of conditions that may be uncommon in contemporary riverscapes. The quandary is well known that ATs (like equilibrium measures of gene flow) have the most power for detecting dispersal when dispersal is least frequent

(Waples and Gaggiotti 2006). When dispersal is common, population allele frequencies are too similar to discriminate immigrants from residents. ATs thus have the greatest utility for detecting dispersal when a) populations are well differentiated, but b) dispersers are abundant enough to be detected and sampled. However, in many contemporary watersheds, connectivity is "all-or-nothing": populations either are completely isolated by impassible barriers to movement, or, in the absence of such barriers, strongly connected via extensive gene flow (e.g., Koizumi et al. 2006; Whiteley et al. 2006; Beneteau et al. 2009; see also Chapters 1 and 4). Even when differentiation is weak, ATs may still be useful for studying asymmetries in gene flow (e.g., Hänfling and Weetman 2006) or patterns of IBD (e.g., Castric and Bernatchez 2004), although there is yet no analytical model that relates spatial trends of cross-assignment to demographic parameters. In contrast, the IBD-slope method applied to differentiation statistics is an underutilized technique that could be quite useful for estimating the lifetime dispersal distance of stream fishes. I predict that in the near future, combined use of "demographic" methods such as CMR with genetic-based estimators of dispersal will dramatically enhance our understanding of the spatial ecology of stream fishes. This study provides examples of the types of insights that can be gained, as well as considerations about the appropriateness of the various genetic tools that are available.

ACKNOWLEDGMENTS

REFERENCES

Albanese, B., P.L. Angermeier, and C. Gowan. 2003. Designing mark-recapture studies to reduce effects of distance weighting on movement distance distributions of stream fishes. Transactions of the American Fisheries Society 132:925-939.

Beneteau, C.L., N.E. Mandrak, and D.D. Heath. 2009. The effects of river barriers and range expansion on the population genetic structure and stability in greenside darter (*Etheostoma blennioides*) populations. Conservation Genetics 10:477-487.

Burkhead, N.M. 1983. Ecological studies of two potentially threatened fishes (the orangefin madtom, *Noturus gilberti* and the Roanoke logperch, *Percina rex*) endemic to the Roanoke River drainage. Final Report to the U.S. Army Corps of Engineers, Wilmington, NC.

Caballero, A. 1994. Developments in the prediction of effective population size. Heredity 73:657-679.

Castric, V., and L. Bernatchez. 2004. Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase of genetic differences with distance. Molecular Ecology 13:1299-1312.

Danancher, D., J.I. Izquierdo, and E. Garcia-Vasquez. 2008. Microsatellite analysis of relatedness structure in young of the year of the endangered *Zingel asper* (Percidae) and implications for conservation. Freshwater Biology 53:546-557.

Dutton, D.J., J.H. Roberts, P.L. Angermeier, and E.M. Hallerman. 2008. Microsatellite markers for endangered Roanoke logperch (*Percina rex*), and their potential utility for other species. Molecular Ecology Resources 8:831-834.

Ensign, W.E., K.N. Leftwich, P.L. Angermeier, and C.A. Dolloff. 1997. Factors influencing stream fish recovery following a large-scale disturbance. Transactions of the American Fisheries Society 126: 895-907.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin version 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Fausch, K.D., C.E. Torgersen, C.V. Baxter, and H.W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52:483-498.

Frankham, R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genetical Research 66:95-107.

Fraser, D.F. and T.E. Sise. 1980. Observations on stream minnows in a patchy environment: a test of a theory of habitat distribution. Ecology 61:790-797.

Gagen, C.J., R.W. Standage, and J.N. Stoeckel. 1998. Ouachita madtom (*Noturus lachneri*) metapopulation dynamics in intermittent Ouachita mountain streams. Copeia 1998:874-882.

Gerking, S.D. 1953. Evidence for the concepts of home range and territory in stream fishes. Ecology 34:347-365.

Goudet, J. 2002. FSTAT: a program to estimate and test gene diversity and fixation indices. Available online at http://www.unil.ch/izea/softwares/fstat.html

Gowan, C., M.K. Young, K.D. Fausch, and S.C. Riley. 1994. Restricted movement in resident stream salmonids: a paradigm lost? Canadian Journal of Fisheries and Aquatic Sciences 51:2626-2637.

Greenberg, L.A. 1991. Habitat use and feeding behavior of thirteen species of benthic stream fishes. Environmental Biology of Fishes 31:389-401.

Greenwood, P.J. 1980. Mating systems, philopatry, and dispersal in birds and mammals. Animal Behavior 28:1140-1162.

Hänfling, B., and D. Weetman. 2006. Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the river sculpin, *Cottus gobio*. Genetics 173:1487-1501.

Hall, C.A.S. 1972. Migration and metabolism in a temperate stream ecosystem. Ecology 53:585-604.

Hall, L.A., and 10 co-authors. 2009. Characterizing dispersal patterns in a threatened seabird with limited genetic structure. Molecular Ecology 18:5074-5085.

Hardy, O.J., and X. Vekemans. 2002. SPAGeDi: a versatile computer program to analyze spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2:618-620.

Harvey, B.C. 1987. Susceptibility of young-of-the-year fishes to downstream displacement by flooding. Transactions of the American Fisheries Society 116:851-855.

Hauser, L., T.R. Seamons, M. Dauer, K.A. Naish, and T.P. Quinn. 2006. An empirical verification of population assignment methods by marking and parentage data: hatchery and wild steelhead (*Oncorhynchus mykiss*) in Forks Creek, Washington, USA. Molecular Ecology 15:3157-3173.

Hedrick, P.W. 2009. Genetics of Populations, 4th Edition. Jones and Bartlett Publishers, Sudbury, Massachusetts.

Helfman, G.S., B.B. Collette, D.E. Facey, and B.W. Bowen. 2009. The Diversity of Fishes: Biology, Evolution, and Ecology, 2nd Edition. Wiley-Blackwell, Oxford.

Hendry, A.P., V. Castric, M.T. Kinnison, and T.P. Quinn. 2004. The evolution of philopatry and dispersal: homing versus straying in salmonids. In Evolution Illuminated: Salmon and Their Relatives (A.P. Hendry and S.C. Stearns, eds.), pp. 52–91. Oxford: Oxford University Press.

Hill, M.F., A. Hastings, and L.W. Botsford. 2002. The effects of small dispersal rates on extinction times in structured metapopulation models. The American Naturalist 160:389-402.

Howell, D.F. 2000. Resampling Procedures, version 1.3. Available online at: www.uvm.edu/~dhowell/StatPages/StatHomePage.html.

Hudy, M., J.A. Coombs, K.H. Nislow, and B.H. Letcher. 2010. Dispersal and within-stream spatial population structure of brook trout revealed by sibship and parentage analysis. Transactions of the American Fisheries Society 139:1276-1287.

Humston, R., B.M. Priest, W.C. Hamilton, and P.E. Bugas, Jr. 2009. Dispersal between tributary and main-stem rivers by juvenile smallmouth bass evaluated using otolith microchemistry. Transactions of the American Fisheries Society 139:171-184.

Jenkins, R.E., and N.M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Jones, O.R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10:551-555.

Kanno, Y., J.C. Vokoun, and B.H. Letcher. 2011. Sibship reconstruction for inferring mating systems, dispersal and effective population size in headwater brook trout (*Salvelinus fontinalis*) populations. Conservation Genetics 12:619-628.

Koizumi, I., S. Yamamoto, and K. Maekawa. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden char. Molecular Ecology 15:3175-3189.

Labbe, T.R., and K.D. Fausch. 2000. Dynamics of intermittent stream habitat regulate persistence of a threatened fish at multiple scales. Ecological Applications 10:1774-1791.

Labonne, J., and P. Gaudin. 2005. Exploring population dynamics patterns in a rare fish, *Zingel asper*, through capture-mark-recapture methods. Conservation Biology 19:463-472.

Leblois, R., A. Estoup, and F. Rousset. 2003. Influence of mutational and sampling factors on the estimation of demographic parameters in a "continuous" population under isolation by distance. Molecular Biology and Evolution 20:491-502.

Leblois, R., F. Rousset, and A. Estoup. 2004. Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. Genetics 166:1081-1092.

Lin, J., T.P. Quinn, R. Hilborn, and L. Hauser. 2008. Fine-scale differentiation between sockeye salmon ecotypes and the effect of phenotype on straying. Heredity 101:341-350.

Lowe, W.H., and F.W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051.

Luikart, G., N. Ryman, D.A. Tallmon, M.K. Schwartz, and F.W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conservation Genetics 11:355-373.

Manel, W., O.E. Gaggiotti, and R.S. Waples. 2005. Assignment methods: matching biological questions with appropriate techniques. Trends in Ecology and Evolution 20:136-142.

Morrissey, M.B., and M.M. Ferguson. 2011. Individual variation in movement throughout the life cycle of a stream-dwelling salmonid fish. Molecular Ecology 20:235-248.

Nathan, R., G. Perry, J.T. Cronin, A.E. Strand, and M.L. Cain. 2003. Methods for estimating long-distance dispersal. Oikos 103:261-273.

Nosil, P., T.H. Vines, and D.J. Funk. 2005. Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution 59:705-719.

Northcote, T.G. 1978. Migratory strategies and production in freshwater fishes. Pages 326-359 in S. D. Gerking, ed. Ecology of Freshwater Fish Production. Blackwell Scientific, Oxford, England

Paetkau, D., L. Waits, P.L. Clarkson, L. Craighead, and C. Strobeck. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genetics 147:1943-1957.

Paetkau, D., R. Slade, M. Burden, and A. Estoup. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. Molecular Ecology 13:55-65.

Page, L.M. 1983. Handbook of Darters. Tropical Fish Hobbyist Publications, Neptune City, NJ.

Palstra, F.P., and D.E. Ruzzante. 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? Molecular Ecology 17:3428-3447.

Peakall, R. and P.E. Smouse. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.

Piry, S., A. Alapetite, J.M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GeneClass2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95:536-539.

Quinn, T.P. 1993. A review of homing and straying of wild and hatchery-produced salmon. Fisheries Research 18:29-44.

Raeymaekers, J.A.M., D. Raeymaekers, I. Koizumi, S. Geldof, and F.A.M. Volckaert. 2009. Guidelines for restoring connectivity around water mills: a population genetic approach to the management of riverine fish. Journal of Applied Ecology 46:562-571.

Rannala, B., and J.L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences 94:9197-9221.

Ridgway, M.S. and B.J. Shuter. 1996. Effects of displacement on the seasonal movements and home range characteristics of smallmouth bass in Lake Opeongo. North American Journal of Fisheries Management 16:371-377.

Roberts, J.H., and P.L. Angermeier. 2007. Spatiotemporal variability of stream habitat and movement of three species of fish. Oecologia 151:417-430.

Roberts, J.H., and P.L. Angermeier. 2010. Assessing impacts of the Roanoke River Flood Reduction Project on the endangered Roanoke logperch. Final Report to the U.S. Army Corps of Engineers, Wilmington, NC.

Roberts, J.H., and A.E. Rosenberger. 2008. Threatened fishes of the world: *Percina rex* (Jordan and Evermann 1889) (Percidae). Environmental Biology of Fishes 83:439-440.

Roberts, J.H., A.E. Rosenberger, B.W. Albanese, and P.L. Angermeier. 2008. Movement patterns of endangered Roanoke logperch (*Percina rex*). Ecology of Freshwater Fish 17:374-381.

Rodriguez, M.A. 2002. Restricted movement in stream fish: the paradigm is incomplete, not lost. Ecology 83:1-13.

Ronce, O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. Annual Review of Ecology and Systematics 38:231-253.

Rosenberger, A.E., and P.L. Angermeier. 2003. Ontogenetic shifts in habitat use by the endangered Roanoke logperch (*Percina rex*). Freshwater Biology 48:1563-1577.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics 145:1219-1228.

Rousset, F. 2000. Genetic differentiation between individuals. Journal of Evolutionary Biology 13:58-62.

Schlosser, I.J. 1987. A conceptual framework for fish communities in small warmwater streams. Pages 17-24 in W. J. Matthews and D. C. Heins, eds. Community and Evolutionary Ecology of North American Stream Fishes. Oklahoma University Press, Norman.

Schlosser, I.J. 1995a. Critical landscape attributes that influence fish population dynamics in headwater streams. Hydrobiologia 303:71-81.

Schlosser, I.J. 1995b. Dispersal, boundary processes, and trophic-level interactions in streams adjacent to beaver ponds. Ecology 76:908-925.

Schwalb, A.N., M.S. Poos, and J.D. Ackerman. 2010. Movement of logperch – the obligate host fish for endangered snuffbox mussels: implications to mussel dispersal. Aquatic Sciences 73:223-231.

Slack, W.T., S.T. Ross, and J.A. Ewing. 2004. Ecology and population structure of the bayou darter, *Etheostoma rubrum*: disjunct riffle habitats and downstream transport of larvae. Environmental Biology of Fishes 71:151-164.

Takahata, N., and M. Nei. 1984. F_{ST} and G_{ST} statistics in the finite island model. Genetics 107:501-504.

Tallmon, D.A., A. Koyuk, G. Luikart, and M.A. Beaumont. 2008. ONeSAMP: A program to estimate effective population size using approximate Bayesian computation. Molecular Ecology Resources 8:299-301.

Taylor, C.M., and M.L. Warren. 2001. Dynamics in species composition of stream fish assemblages: environmental variability and nested subsets. Ecology 82:2320-2330.

Turner, T.F. 2001. Comparative study of larval transport and gene flow in darters. Copeia 2001:766-774.

Turner, T.F., M.J. Osborne, G.R. Moyer, M.A. Benavides, and D. Alo. 2006. Life history and environmental variation interact to determine effective population to census size ratio. Proceedings of the the Royal Society B 273:3065-3073.

VanArnum, C.J.G., G.L. Buynak, and J.R. Ross. 2004. Movement of smallmouth bass in Elkhorn Creek, Kentucky. North American Journal of Fisheries Management 24:311-315.

Van Oosterhout, C., W.F. Hutchinson, D.P.M. Wills, and P. Shipley. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

Vera, M., N. Sanz, M.M. Hansen, A. Almodóvar, and J.L. Garcia-Marín. 2010. Population and family structure of brown trout, *Salmo trutta*, in a Mediterranean stream. Marine and Freshwater Research 61:672-681.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. Genetics 166:1963-1979.

Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology 18:2148-2164.

Wang, J., and A.W. Santure. 2009. Parentage and sibship inference from multi-locus genotype data under polygamy. Genetics 181:1579-1594.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics 7:167-184.

Waples, R.S., and C. Do. 2008. LDNe: A program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753-756.

Waples, R.S., and C. Do. 2010. Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications 3:244-262.

Waples, R.S., and O. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419-1439.

Waser, P.M., and C. Strobeck. 1998. Genetic signatures of interpopulation dispersal. Trends in Ecology and Evolution 13:43-44.

Weir, B.S., and C.C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.

Whiteley, A.R., P. Spruell, and F.W. Allendorf. 2006. Can common species provide valuable information for conservation? Molecular Ecology 15:2767-2786.

Whitlock, M.C., and D.E. McCauley. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. Heredity 82:117-125.

Winemiller, K.O., and K.A. Rose. 1992. Patterns of life-history diversification in North American fishes: implications for population regulation. Canadian Journal of Fisheries and Aquatic Sciences 49:2196-2218.

Winn, H. E. 1958. Comparative reproductive behavior and ecology of fourteen species of darters (Pisces-Percidae). Ecological Monographs 28:155-191.

Wright, S. 1943. Isolation by distance. Genetics 28:114-138.

Table 2.1. DNA sample sizes for juvenile, candidate father, and candidate mother *Percina rex*, by sampling site and year. Site locations are shown in Figure 2.1. Dashes indicate sites not sampled in 2005.

Site	Juvenile cohort 2005	Juvenile cohort 2006	Candidate fathers 2005	Candidate fathers 2006	Candidate mothers 2005	Candidate mothers 2006
RR1	0	2	0	0	1	0
RR2	0	6	0	0	0	0
RR3	4	9	0	0	2	0
RR4	5	3	0	1	1	0
RR5	1	3	0	3	0	1
RR6	0	8	0	0	1	0
RR7	0	3	0	1	0	1
RR8	10	6	0	2	2	2
RR9	4	6	0	0	0	1
RR10	1	5	0	2	3	1
RR11	0	5	0	2	0	2
RR12	-	1	-	0	-	0
SF	-	9	-	2	-	0
NF1	-	1	-	0	-	0
NF2	-	2	-	4	-	1
Total	25	69	0	17	10	9

Table 2.2. Mean and bootstrapped 95% confidence interval (CI) of the slope of the relationship between the genetic (F_{ST} or $\hat{a}$) and spatial distance separating pairs of sites or individuals, over selected distance ranges.

Response	Distance range (km)	Pair-wise Comparisons	Isolation-by-distance slope (β)	
			Mean	95% CI
F_{ST}	1 - 55 (full)	78	0.00025	(-0.00010, 0.00060)
F_{ST}	5 - 55	62	0.00024	(-0.00015, 0.00063)
F_{ST}	10 - 55	45	0.00027	(-0.00029, 0.00083)
F_{ST}	1 - 40	65	0.00056	(-0.00004, 0.00116)
F_{ST}	1 - 25	55	-0.00057	(-0.00168, 0.00054)
$\hat{a}$	1 - 55 (full)	3985	0.00029	(-0.00047, 0.00105)
$\hat{a}$	5 - 55	3000	0.00029	(-0.00049, 0.00107)
$\hat{a}$	10 - 55	2288	0.00036	(-0.00042, 0.00114)
$\hat{a}$	1 - 40	3499	0.00034	(-0.00071, 0.00139)
$\hat{a}$	1 - 25	2957	0.00050	(-0.00081, 0.00181)
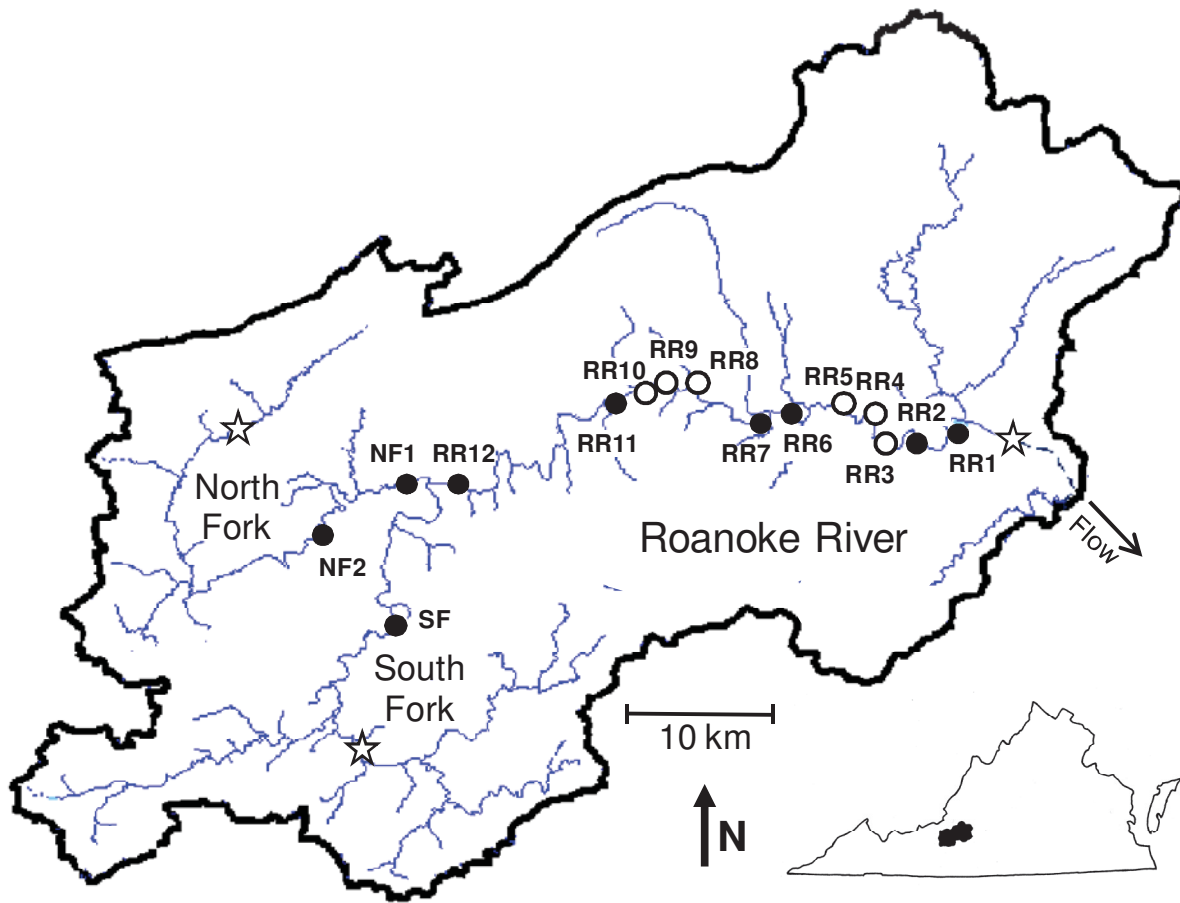
Figure 2.1. The upper Roanoke River watershed (upstream of Smith Mountain Reservoir) in Virginia, USA. Juvenile *Percina rex* were captured at six sites in 2005 (open circles) and at these plus an additional nine sites in 2006 (filled circles; see Table 1). Approximate distributional limits of *P. rex* within the watershed are indicated by stars.
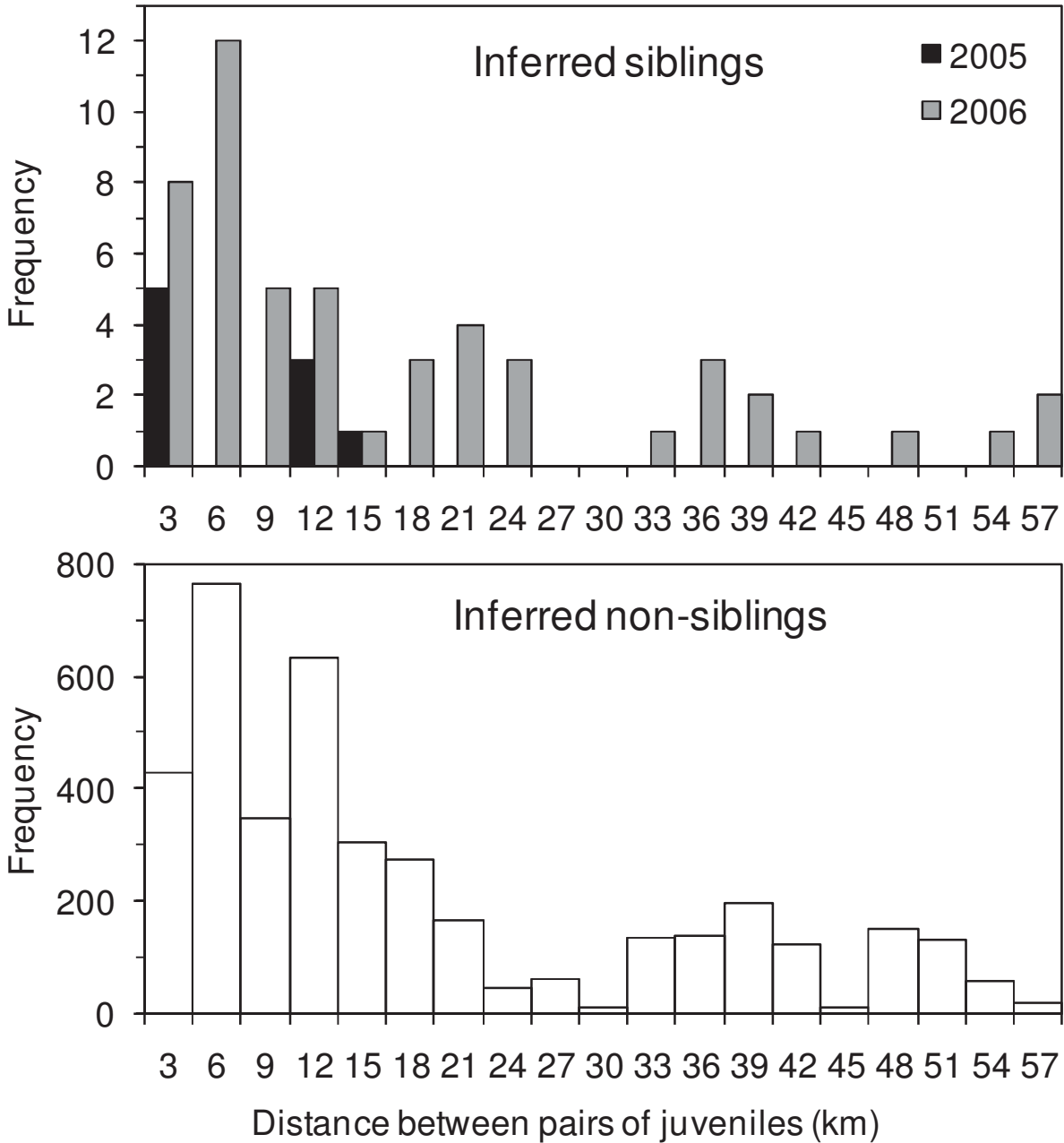
Figure 2.2. Frequency histograms of distances separating capture locations of half-sibling and non-sibling juvenile *Percina rex*, as deduced through genetic pedigree reconstruction. Black bars refer to pairs of siblings both captured in 2005, grey bars to pairs in which at least one member was captured in 2006. Maximum detectable spatial separation was 18 km in 2005 and 55 km in 2006 (see text). Neither the mean nor shape of distributions differed between siblings and non-siblings (both P > 0.05).
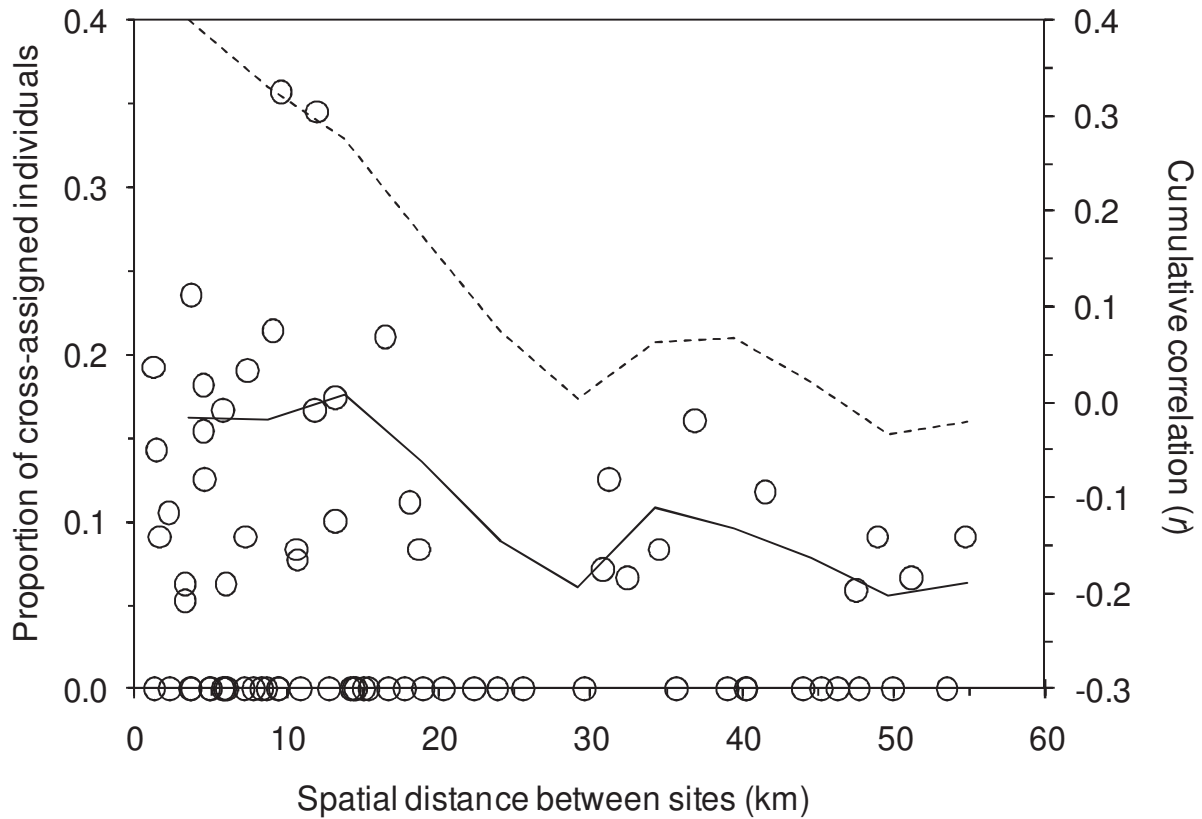
Figure 2.3. Variation in the proportion of juvenile *Percina rex* assigned by assignment tests (ATs) to a site other than the one at which they were captured, as a function of the spatial distance between sites. Each open circle compares one pair of sites. The solid and dashed lines indicate the mean and upper 95% bootstrapped confidence limit of the correlation (r) between distance and cross-assignment, at cumulatively increasing 5-km increments (i.e., 1-5 km, 1-10 km, etc.). Confidence limits overlap zero over all distance increments except 1-50 km and 1-55 km.
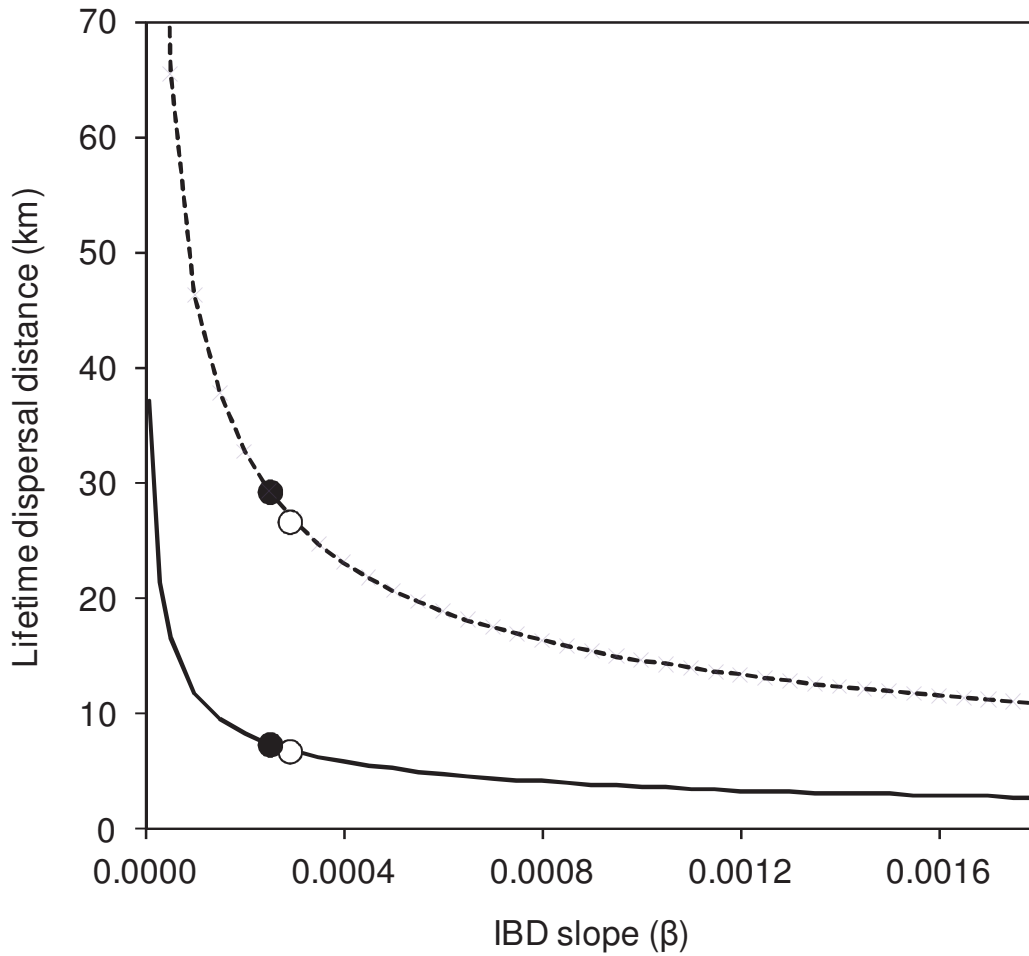
Figure 2.4. Variation in predicted mean lifetime dispersal distance of *Percina rex* (see Equation 1) over the range of uncertainty around the isolation-by-distance slope (β) and the effective breeder density (*D*). Based on the range of estimates of effective population size, confidence limits of *D* ranged from 1.2 (dotted line) to 18.2 (solid line) breeders km^{-1}. Mean estimates of β from group- and individual-based differentiation were 0.00025 (filled circles) and 0.00029 (open circles), respectively. Confidence limits for β ranged from -0.00168 to 0.00181, but only the positive portion of this range is shown.

CHAPTER 3: Designing risk-averse conservation strategies for fragmented stream-fish populations

ABSTRACT

Many populations of stream fish persist in small remnant habitat patches that are fragmented by anthropogenic barriers to dispersal. Conservation of such populations often involves one of two very different strategies: (1) restoration of connectivity, which seeks to minimize risks associated with small population size (e.g., demographic and environmental stochasticity, inbreeding depression, and genetic drift), or (2) maintenance of isolation, which seeks to minimize risks associated with the mixing of historically separated gene pools (e.g., introgression, outbreeding depression, and loss of local adaptations). The optimal strategy for each population depends on the relative magnitudes of these "small-population" (SP) versus "outbreeding-depression" (OD) risks, which in turn depend upon the contemporary distribution, abundance, and genetic diversity of populations and historical evolutionary relationships between populations. I developed a risk-assessment framework to facilitate the joint estimation of SP and OD risks based on published risk criteria and estimable risk metrics. To illustrate how the framework could be used to evaluate risk and guide conservation choices, I applied it to populations of Roanoke logperch (*Percina rex*), an endangered riverine fish whose distribution is highly fragmented. Risk metrics were calculated from attributes of *P. rex*'s population density and geographic distribution and from genetic parameters estimated from nuclear and mitochondrial DNA markers. Both SP and OD risks varied considerably among populations, suggesting that a blanket management policy would be ineffective. Only one of seven populations exhibited low risk across both types, whereas four populations exhibited moderate to high risk for SP (but not OD) and two exhibited moderate risk for OD (but not SP). This information clarifies management strategies for *P. rex*: four populations could benefit from restored connectivity, two would be better protected in isolation, and one tentatively is secure. These management options should be re-evaluated periodically, whenever new data on population status become available. Such a risk-assessment approach can be useful for developing scientifically defensible conservation strategies for *P. rex*, as well as for other threatened fishes.

INTRODUCTION

Over the past few centuries, humans have dramatically accelerated the extinction rate of other species (Pimm et al. 1995). Species rarely go extinct all at once; most disappear gradually via the incremental extirpation of constituent populations (Caughley 1994). As such, populations are often the foci of conservation intervention. Populations also constitute cohesive demographic and evolutionary units, and generally occur over spatial extents that are feasible to manage (McElhany et al. 2000). The persistence and evolutionary potential of many populations are threatened by a host of factors, including habitat loss and fragmentation, overharvest, disease, introduced species, demographic and environmental stochasticity, inbreeding and outbreeding depression, and the loss of adaptive potential due to genetic drift (Groom et al. 2006). For some populations, these risk factors can be estimated and ranked (Allendorf et al. 1997). Mitigating the impacts of drivers of extirpation requires an optimization routine in which 1) the biological risks associated with different management options are quantified, and 2) those options that minimize risk are pursued.

Stream fishes are among the most imperiled groups of North American fauna, due primarily to hydrologic and landscape alteration by humans (Ricciardi and Rasmussen 1999). Stream fishes face the typical range of threats described above, but are particularly vulnerable to habitat fragmentation via the construction of anthropogenic barriers to dispersal (Fagan 2002). As a result, many contemporary populations of stream fish persist in small remnant habitat patches that are demographically and genetically isolated from conspecifics. Small, isolated populations face an increased risk of extinction due to elevated demographic and environmental stochasticity, inbreeding depression, genetic load, and lack of demographic rescue from immigrants (Morita and Yamamoto 2002; Winston et al. 2001). They also face an increased risk of inadequate long-term evolutionary potential, due to the rapid loss of genetic diversity to drift in small populations (Allendorf and Leary 1986; Pritchard et al. 2007; Skalski et al. 2008). Collectively, investigation of these "small-population risks" (henceforth "SP risks") has formed much of the basis of the discipline of conservation biology (Caughley 1994).

A variety of management strategies are available for minimizing SP risks, but not all are feasible in all situations. For example, curtailing species harvest, restoring habitat quality, or expanding range extent theoretically can increase population size and thereby decrease SP risks. However, many threatened fish species already are protected from harvest and/or are not the focus of fisheries. Additionally, the habitat features limiting suitability may be unknown or impossible to restore. Similarly, range expansion may be impossible for a species that is limited by stream size or other natural habitat features, or infeasible if, for example, it would require the draining of a reservoir. In many cases, the only remaining strategy left to managers for boosting population size and genetic diversity may be the restoration of connectivity *between* populations. This could involve the restoration of natural dispersal via the removal of barriers or the initiation of managed dispersal via the translocation of fish from captive or other wild populations. Such conservation strategies have been used successfully to reduce SP risks in various taxa (Hedrick and Fredrickson 2010), including fishes (Vrijoenhoek 1994; Minckley et al. 2003; Yamamoto et al. 2006).

Although restoring connectivity may decrease SP risks, it may increase other types of risk. Increased connectivity can facilitate the immigration of undesirable immigrants such as novel pathogens and invasive species (Dunham et al. 2002; Smith and Jones 2005; Fausch et al. 2009). Furthermore, connecting formerly disconnected gene pools can have unintended negative fitness consequences for offspring, including outbreeding depression and the loss of local adaptations due to genetic swamping by immigrants (Tallmon et al. 2004; Edmands 2007). The likelihood of these "outbreeding-depression risks" (henceforth "OD risks") is greatest when fish are translocated among populations that are strongly evolutionarily diverged. This divergence is best assessed using adaptively significant trait variation, but because such data rarely are available for wild populations, often is assessed using proxy measures of geographical, environmental, phenotypic, or neutral genetic divergence (Edmands 2002). In general, OD risks are more difficult than SP risks to quantify and have been less investigated than SP risks (Edmands 2007). Nevertheless, as aquatic fauna become increasingly fragmented and imperiled, questions about whether to restore connectivity or maintain isolation will become more frequent and contentious

(Minckley et al. 2003; George et al. 2009; Fausch et al. 2009). Conservation biologists need transparent and scientifically defensible decision-making tools to weigh the relative risks and to select an optimal conservation strategy (Francis and Shotton 1997).

A framework for assessing risks to fragmented fish populations

I developed a risk-assessment framework to facilitate the joint estimation of SP and OD risks to fragmented fish populations based on published risk criteria and estimable risk metrics. Although similar tools are available for assessing SP (Allendorf et al. 1997; Lindley et al. 2007; Peterson et al. 2008) or OD (Emlen 1991; Frankham et al. 2011) risks to populations, I know of no previous attempt to unite these two risk types into a common framework. Fausch et al. (2009; see also Peterson et al. 2008) have developed a conceptually similar framework to guide decisions about restoring connectivity versus maintaining isolation of westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) populations. However, their framework differs from mine in that they focus on ecological risks of inter-specific interactions following barrier removal, whereas I focus on intra-specific genetic interactions following translocations.

My framework employs a series of individual metrics, categorizing metric values into "low", "moderate", or "high" risk scores based on a set of risk criteria (Table 3.1). Metric scores then are averaged within the SP and OD categories to develop an aggregate risk score for each risk type. My expectation is that the resulting ordination of populations in SP-OD "risk space" will lead to prescriptions about proper management options (Figure 3.1). If one type of risk predominates over another, the optimal management strategy should be clear and uncontroversial. For example, a population with moderate to high SP risk but low OD risk needs restored connectivity, whereas a population with moderate to high OD risk but low SP risk should be maintained in isolation. In contrast, when populations are at high risk for both risk types, the optimal management strategy is unclear, because any action will involve trading off one form of risk for another. Such situations are likely to be scientifically contentious and ultimately resolved by non-scientific value judgments about which evolutionary processes are most deserving of conservation (e.g., Hedrick 1995).

83

A variety of metrics could be employed to gauge SP and OD risks (Allendorf et al. 1997; Edmands 2002; Coltman and Slate 2003; Reed et al. 2003; Frankham et al. 2011); my list contains a representative selection of these (Table 3.1). Only in rare situations would the information necessary to calculate all of these metrics be available (Allendorf et al. 1997). I therefore focused on those metrics I perceived to be most commonly available for threatened stream fishes and particularly those available for Roanoke logperch, the case-study species (see below). Furthermore, although my framework contains specific criteria for distinguishing low-, moderate-, and high-risk situations, my goal is not to promote any particular set of criteria for general use. Rather, I envision the framework as a flexible template within which users can customize metrics and criteria based on the biology of their species, available data types, and advances in our understanding of how risk factors affect persistence and evolution. My primary goal is to make the process of selecting and ranking risks a transparent and repeatable scientific endeavor.

Metrics for measuring SP risks are designed to gauge threats to a population's persistence (e.g., from demographic stochasticity, environmental stochasticity, catastrophes, and inbreeding depression) and long-term adaptive potential (e.g., susceptibility to genetic drift) (Figure 3.1). Population viability analysis (PVA) models are among the most informative predictors of such risks (Reed et al. 2003), but PVA has yet to be attempted for many fishes of conservation concern. Alternative, potentially informative metrics include estimates of census (N) and effective (N_e) population size, and the level of individual inbreeding (as measured by the inbreeding coefficient F). Estimates of the minimum N required for persistence vary between studies from ~100 (Berger 1990) to over 7000 individuals (Reed et al. 2003). Likewise, estimates of the minimum long-term N_e necessary to maintain adaptive potential range from approximately 400-500 (Franklin 1980; Waples 1990) to over 5000 breeding individuals (Franklin and Frankham 1998), with at least 50 breeding individuals required over the short term to avoid inbreeding depression (Franklin 1980). As little as a 10% increase in F can cause population growth rate to decline (McCauley and Wade 1981), and F values of 25% or more (i.e., corresponding to full-sib

84

mating) consistently result in reduced individual fitness and reduced population growth (Edmands 2007; Frankham et al. 2010)

Spatial characteristics of a population's distribution also may be predictive of its susceptibility to SP risks. Range extent correlates positively with persistence for some salmonids (Hilderbrand and Kershner 2000; Young and Harig 2001), but because fish density varies idiosyncratically, quantitative range-extent criteria likely would not be transferable between species. On the other hand, range complexity may be a more transferable metric. For a given abundance, environmental stochasticity is less likely to cause extirpation if a population occurs over a greater diversity of habitats that are environmentally uncorrelated (McElhany et al. 2000). Instream disturbances, for example, are unlikely to affect all portions of a stream network equally, so a distribution that spans more network branches (i.e., stream channels) should increase the likelihood of at least some individuals surviving the disturbance (Townsend et al. 1997; Taylor and Warren 2001; Fagan 2002). Specifically, I expect populations occupying a single stream channel to be at highest SP risk, whereas I expect populations occupying larger numbers of stream channels to experience correspondingly lower SP risk.

Although the best indicators of OD risk would measure divergence in functional genes and genetically-based traits that influence fitness (Emlen 1991; Crandall et al. 2000), such data are absent for many species of conservation concern (Frankham et al. 2011). Measures of geographical, environmental, phenotypic, or neutral genetic divergence can be used to estimate divergence time and the potential for differential selection pressures, two surrogates for OD risk (Edmands 2002). Divergence time relates positively to OD risk, though the relationship is nonlinear and varies among taxa. In Edmands's (2002) review of intra- and inter-specific crosses in various vertebrate and invertebrate taxa, OD first emerged after anywhere from 10,000 to 350,000 years of isolation, but reproductive compatibility persisted for 8 to 56 million years. In a more recent meta-analysis of intra-specific crosses, Frankham et al. (2011) found that >20 generations of isolation in selectively different environments or >500 years of isolation in selectively similar environments were required before OD was observed. Genetic distance measures like G_{ST} or F_{ST} generally show no clear relationship to OD risks (McClelland and Naish 2006; Edmands 2007;

but see Fraser et al. 2010), whereas the percentage divergence at DNA sequences relates positively to inter-specific reproductive compatibility in fishes (Russell 2003; Bolnick and Near 2005), suggesting a possible relationship to intra-specific OD risk. Again, however, generalities are difficult to draw. In some studies, reduced offspring viability began after as little as 0.2% mitochondrial DNA (mtDNA) sequence divergence between parents (Edmands 2007), whereas in another study, OD was not observed until parents exhibited at least 7% sequence divergence (Russell 2003).

Potential measures of geographical and environmental divergence include the fluvial distance separating populations and the biogeographic history and physiographic features experienced by populations. OD risks should increase with fluvial distance, because larger distances potentially span a wider range of environmental conditions (Angermeier and Winston 1999) and correspond to lower gene flow (assuming genetic isolation-by-distance; see Chapter 1). However, the slopes of these relationships probably are context specific, due to variability in species' dispersal patterns and landscape heterogeneity. Geographic strata such as stream-basin boundaries and physiographic provinces explain much of the variation in fish-assemblage composition among streams (Angermeier and Winston 1999), suggesting that these strata might also capture distinct evolutionary processes of significance for assessing OD risks.

Case study: application to endangered Roanoke logperch

To illustrate how the framework could be used to guide conservation choices, I applied it to populations of Roanoke logperch (*Percina rex*), an endangered riverine fish. *P. rex* is a large-bodied darter (Teleostei: Percidae) restricted to streams and rivers in the Roanoke, Dan, and Nottoway river basins of Virginia and North Carolina, USA (Roberts and Rosenberger 2008). The seven known extant populations are demographically isolated by large dams and reservoirs and long reaches of unsuitable habitat (see Chapter 1) and presumably vary widely in size and vulnerability to human activities (Rosenberger 2007). The species' most recent recovery plan (Rosenberger 2007) outlines the need to assess and increase the viability of populations and restore connectivity among populations. However, there exist no estimates of demographic trends with which to assess the viability outlook for *P. rex* populations. Furthermore, although it has been hypothesized that presently disjunct populations were

more connected prior to European settlement of Virginia (Jenkins and Burkhead 1994), and thus that the risks of mixing populations might be slight, this hypothesis has never been tested with empirical data. Such uncertainty complicates the formation of defensible conservation strategies for this species.

I used available demographic, geographic, and genetic information to conduct an assessment of SP and OD risks to populations of *P. rex*. For each risk type, I estimated risk level for a suite of metrics, as well as an aggregate risk score across metrics. Results showed clear separation of populations in "risk space", suggesting that a one-size-fits-all management strategy would be ineffective for this species. Understanding these relative risks provides clear guidance on how to tailor management strategies to individual populations of *P. rex*. Such a risk-ranking system could be extended to other fragmented fish populations.

METHODS

Genetic data

I previously analyzed variation of 11 nuclear DNA microsatellite loci in 578 individual *P. rex* and delineated seven completely isolated, yet internally panmictic populations: upper Roanoke River and tributaries (UROAN), Pigg River and tributaries (PIGG), Goose Creek (GOOSE), Big Otter River and tributaries (OTTER), Smith River downstream of Philpott Reservoir (LSMITH), Smith River upstream of Philpott Reservoir (USMITH), and Nottoway River and tributaries (NOTT; see Chapter 1). Herein, I selected a random sample of 30 individuals from each of these populations for further study. For these 210 individuals, I analyzed variation across 1037 bp of the *ND2* mitochondrial DNA (mtDNA) gene and reanalyzed variation at the 11 microsatellite loci. These data were used to estimate six standard measures of genetic diversity within populations (haplotype and nucleotide diversity, number of segregating sites, and number of private haplotypes for *ND2*; unbiased gene diversity and allele richness for microsatellites) and four standard measures of genetic differentiation between populations [percent between-population sequence divergence and Hudson et al.'s (1992) F_{ST} for *ND2*; Weir and Cockerham's (1984) F_{ST} and Slatkin's (1995) R_{ST} for microsatellites]. I estimated microsatellite statistics using Arlequin 3.11 (Excoffier et al. 2005) or Populations 1.2.3 (O. Langella; http://bioinformatics.org/~tryphon/populations),

and I estimated *ND2* statistics using DNAsp 5.1 (Librado and Rozas 2009) or MEGA 5.05 (http://www.megasoftware.net/). These statistics were used in the assessment of risk, as well as during the estimation of demographic parameters in simulation models (see below). More detailed summaries of the microsatellite and *ND2* studies are provided in Chapter 1 and Appendices D and E.

I used approximate Bayesian computation (ABC) models to estimate demographic parameters of interest for assessing risk. ABC models are useful for gaining insight into demographic histories too complex for genetic likelihood-based methods to accommodate, for example histories involving fragmentation and bottleneck events (Beaumont et al. 2002; Lopes and Boessenkool 2010). Details of model development are in Appendix F. Briefly, the principle of ABC is to iteratively simulate millions of demographic scenarios, each iteration drawing parameter estimates (e.g., N_e, m) from plausible prior distributions. From each simulated dataset, a series of genetic summary statistics is calculated. Following all simulations, Euclidean distances are calculated between simulated statistics and empirical statistics obtained from real populations, and simulations with large distances are rejected. Posterior estimates of demographic parameters are then obtained from retained simulations.

I parameterized demographic simulations based on the presumed history of *P. rex* (see Appendix F for details). The model consisted of seven populations joined in a hierarchical metapopulation structure. Populations exhibited constant historical size and exchanged migrants at a constant historical rate (m) from the time of founding until a point, t generations in the past, at which migration ceased and population size instantaneously declined by a certain percentage (B) to a new, contemporary size (N_e). Each population was assigned its own N_e and B values, whereas pairs of populations were assigned one of four m values, depending on whether the comparison was within basins, between the Roanoke and Dan basins, between the Roanoke and Nottoway basins, or between the Dan and Nottoway basins. The true values of these 19 demographic parameters (i.e., seven Bs, seven N_es, four ms, and one t,) and the mutation rate (μ) were unknown, so I treated them as random variables and assigned them diffuse prior distributions that bracketed the plausible range of values (Appendix F). At each model iteration, a random value was drawn from each parameter's prior distribution and used to parameterize a simulated

demographic history. Simulated data were generated in BayeSSC

(http://www.stanford.edu/group/hadlylab/ssc/index.html) and the ABC rejection step was performed in R

2.10.0 (R Core Development Team) using a script written by C. Anderson (Harvard University). Separate

models were run for microsatellite and *ND2* data. In each case, I performed 5,000,000 simulations,

updating priors every 1,000,000 simulations. Posteriors were estimated from the final 500 accepted

simulations. I assumed a 1:1 ratio of males:females and doubled the female N_e estimated by mtDNA

models to derive an estimate of total N_e. The harmonic mean of this estimate and the N_e estimated from

microsatellite models was employed as a metric for risk assessment (see below).

A probabilistic estimate of the divergence time (*t*) of all populations was estimated from ABC

models described above. To complement this estimate, I developed point estimates of *t* between each pair

of populations using Nei and Chakravarti's (1977) isolation model. The model predicts that F_{ST} increases

over time in completely isolated populations as function of N_e and *t*, approximately as:

$$F_{ST} \approx 1 - e^{(-t / 2Ne)}$$

I rearranged this equation to solve for *t* based on values of F_{ST} and the harmonic mean N_e for pairs of

populations, as estimated from microsatellite data. To convert between *t* in generations and time in years

requires an estimate of *P. rex*'s generation time. The species matures at 2.5 years and lives to 6.5 years

(Jenkins and Burkhead 1994), so I assumed a generation time at the midpoint of this interval, 4.5 years.

I estimated each population's susceptibility to inbreeding depression using the inbreeding

coefficient (*F*). Lacking knowledge of current levels of inbreeding, I estimated the percentage increase of

F after 100 years of further inbreeding at the current N_e, assuming a generation time of 4.5 years

(Frankham et al. 2010). Because populations likely already have experienced some inbreeding (i.e.,

current *F* > 0), this provides a minimum estimate of the total future level of inbreeding.

Geographic and demographic data

I estimated the total geographic extent and number of stream channels occupied by each

population of *P. rex* based on published distributional data (reviewed in Rosenberger 2007) and personal

communication with agency personnel. Once the upstream and downstream distributional limits were

determined for a population, I calculated range extent as the total length of stream (km) between these points, measured in Google Earth 5.1 (http://www.google.com/earth). Data on population density (fish km^{-1}) were extracted from previous studies (see Appendix G for details) and multiplied by range extent to develop coarse estimates of total population size (N) for each population. To index potential OD risks due to environmental differences experienced by populations, I recorded whether a population occupied a unique basin (Roanoke, Dan, or Nottoway) or physiographic province (Valley and Ridge, Piedmont, or Coastal Plain). I also estimated the fluvial (i.e., along the stream) distance from each population to its nearest neighboring population in Google Earth.

Calculation of risk scores

I used a matrix to quantify risks to each population across four SP and four OD risk metrics. Each population was assigned a numerical risk level (1 = low, 3 = moderate, 5 = high) for each metric, based on the metric value and a set of risk criteria (Table 3.1). Scores were then averaged across metrics within risk types and used to ordinate populations in risk space and evaluate management options. My choice of a 1-5 numerical scale was arbitrary; however, the use of alternative three-category scales would not alter the inferred relative risks to populations. SP metrics were calculated from characteristics intrinsic to each population, whereas OD metrics were calculated based on comparisons among populations. In the latter case, I estimated risk based on the *minimum* difference (e.g., genetic, spatial, or environmental distance) between a population and its closest relative, rather than the *average* difference between a population and all other populations. In effect, this sought to answer the question of whether *any* populations were suitable for intermixing, rather than whether *all* populations were suitable intermixing. The eight metrics included: 1) total population size (N), 2) effective population size (N_e), 3) percent increase of F over the next 100 years, 4) number of occupied stream channels, 5) minimum percent sequence divergence from closest relative, 6) fluvial distance to closest neighboring population, 7) minimum divergence time from closest relative, and 8) uniqueness of occupied basin and physiography (Table 3.1).

I adapted numerical risk criteria for N, N_e, F, percent sequence divergence, and divergence time from published studies cited above. Considerable uncertainty surrounded all of these criteria; I generally used the most conservative (i.e., lowest) published cutoffs, preferring to err on the side of caution. Criteria for other metrics were developed based on professional judgment considering the biology of *P. rex*. Presumably, uncertainty over criteria for any given metric was compensated for by the use of multiple metrics, which prevented any single metric from disproportionately driving overall risk score.

Because *P. rex* can maintain panmixia over spatial extents of up to 80 km (see Chapter 1), I presumed that populations within that proximity were unlikely to exhibit adaptively significant divergence and assigned such comparisons a low OD risk score. I arbitrarily set the threshold between medium and high risk at twice this spatial distance. Using a similar rationale, populations sharing the same basin and physiography were presumed to have low OD risk, those sharing one or the other but not both were presumed to have moderate OD risk, and those occurring in a different basin and physiography were presumed to have the highest OD risk.

RESULTS

Based on estimates of range extent and population density (Appendix G), my derived estimates of *P. rex* population size ranged among populations from 719 (in GOOSE) to 13841 (in UROAN) (Table 3.4). Three populations exhibited N large enough (i.e., > 7000) to confer low SP risk, whereas the remaining four exhibited N small enough (i.e., 500-7000) to confer moderate SP risk. The geographic ranges of these populations spanned between one and six stream channels, conferring anywhere from a low to a high SP risk. Three populations co-occurred in the Roanoke basin and Piedmont physiographic province, and two other populations co-occurred in the Dan basin and Piedmont province. Each of these was attributed low risk for the "occupied basin/physiography" metric. In contrast, UROAN occurred in a unique physiography (Valley and Ridge) and NOTT occurred in a unique basin (Nottoway) and physiography (Coastal Plain), so these populations received moderate and high OD risk scores, respectively. The minimum fluvial distance of a population from its nearest neighbor ranged from < 31

km (LSMITH and USMITH) to > 725 km (NOTT). Four populations, two populations, and one population received low, moderate, and high OD risk scores for this metric, respectively.

Estimated levels of genetic diversity varied widely among the seven *P. rex* populations (Table 3.2; Appendices D and E). Population UROAN contained 11 of the 15 total *ND2* haplotypes, and 8 of these were unique to that population. UROAN also exhibited the highest gene diversity and allele richness at microsatellite loci. At the other extreme, populations GOOSE and OTTER exhibited only 1 or 2 *ND2* haplotypes, no unique haplotypes, and the lowest estimated levels of microsatellite gene diversity and allele richness. The only other population to exhibit substantial novel genetic diversity was NOTT, for which both observed haplotypes were unique to that population.

ABC simulation models produced fairly precise estimates of contemporary N_e and the length of time that populations have been isolated, but could not estimate pre-isolation N_e or migration rates with any precision (Appendix F). Point estimates (with 95% credible intervals) of divergence time were 14 (10-67) and 17 (10-37) generations ago based on microsatellite and *ND2* data, respectively (Appendix F). Assuming a generation time of 4.5 years for *P. rex*, this corresponds to 63 (45-302) and 77 (45-167) years ago for microsatellites and *ND2*, respectively, around the time when most dams were constructed in Virginia (Jenkins and Burkhead 1994). Point estimates of contemporary N_e ranged among populations from 11 (2-144) to 1198 (365-3143) breeding individuals per generation based on microsatellites, and from 5 (1-36) to 280 (140-1042) breeding females per generation based on *ND2* (Table 3.2, Appendix F). For both marker types, the estimated effective size of UROAN was substantially larger than any other population, whereas the effective sizes of GOOSE and OTTER were relatively small. The harmonic mean of N_e across both marker types ranged among populations from 26 to 879 breeding individuals per generation. Based on these values, UROAN received a low SP risk score, GOOSE received a high score, and other populations received a moderate score for the N_e risk metric (Table 3.4). Furthermore, assuming these N_e values, the expected increase in inbreeding (*F*) over the next 100 years (i.e., 22 generations) ranged from 1% to 35%. For this metric, GOOSE and OTTER received high and moderate SP risk scores, respectively, and other populations received low risk scores.

Inter-population *ND2* sequence divergence ranged among population pairs from 0.0% to 0.6% (Table 3.3). Most populations were closely related (i.e., divergence 0.0-0.1%) to at least one other population, and thus were assigned low OD risk scores for this metric (Table 3.4). The exception was NOTT, which was 0.3% diverged from its closest relatives (UROAN and USMITH; Table 3.3) and received a moderate risk score. Divergence times estimated from microsatellite data indicated similarly recent isolation of populations. Assuming an isolation model (Nei and Chakravarti 1977), the F_{ST} values observed between pairs of populations were consistent with relatively recent divergence from a common ancestor. The minimum divergence time between a population and its closest relative ranged from 5 generations (20 years; PIGG and GOOSE) to 17 generations (78 years; NOTT). These time-spans are shorter than the thresholds indicative of elevated risk (Table 3.1), so all populations received a low OD risk score for this metric (Table 3.4).

Overall mean risk scores ranged from 1.0 to 4.5 (on a 1 to 5 scale) for SP risks and from 1.0 to 3.5 for OD risks (Table 3.4). Graphical ordination based on risk scores shows clear separation of populations in risk space (Figure 3.2). Only one population (LSMITH) exhibits low risk across both the SP and OD axes, whereas four populations (PIGG, GOOSE, OTTER, USMITH) exhibit moderate to high SP risk but low OD risk, and two populations (UROAN and NOTT) exhibit moderate to high OD risk but low SP risk. Risk types were inversely related, such that no population exhibited elevated risk for both risk types.

DISCUSSION

Considerations in using the risk assessment framework

In this study, I developed a framework to help guide the process of estimating, evaluating, and communicating the risks incurred by adopting different conservation options for fragmented fish populations. Although many of the tradeoffs inherent in isolating versus mixing strategies have been discussed previously (e.g., Crandall et al. 2000; Tallmon et al. 2004; Jones et al. 2006; Edmands 2007; Fausch et al. 2009; George et al. 2009), managers often are left with little specific guidance on how to measure and balance these risks in practice. Faced with this uncertainty, managers may adopt a "do-

nothing" strategy, presuming that this represents the most cautious approach (Francis and Shotton 1997). However, all management options – including the do-nothing strategy – bear some type of risk (Crandall et al. 2000). It therefore is critical that such risks be measured using transparent and repeatable scientific methods, and then explicitly communicated to all interested parties.

In my view, previous management recommendations have inadequately weighed and communicated SP and OD risks to fish populations. This may stem from philosophical differences among investigators in their past experiences and opinions concerning the relative magnitudes of opposing risks (Fausch et al. 2009), real or perceived lack of available data or metrics with which to assess such risks (Edmands 2007), or a combination of the two. Regardless the reason, incomplete and/or subjective interpretation of risk may in the best of cases breed contentiousness and distrust among scientists, managers, and stakeholders, and in the worst of cases lead to improper management of species. My framework can improve this situation because it employs transparent, explicit criteria to develop objective measures of relative susceptibility to both SP and OD risks. Parties still may disagree over which currencies and which criteria are most appropriate for measuring risk in a given species, but such questions can be resolved in an objective, scientific manner.

Past management prescriptions for fragmented fish populations could have benefitted from the use of such a framework. For example, native populations of the endangered watercress darter (*Etheostoma nuchale*) persist in four isolated springs in central Alabama, U.S. (Fluker et al. 2010). These populations have experienced demographic declines and are threatened by ongoing urbanization, low genetic diversity, and small N_es, suggesting moderate to high SP risks. However, most populations also exhibit strong genetic divergence and evidence for long-term isolation, suggesting moderate to high OD risks. *E. nuchale* thus occupies the sector of risk space most challenging to manage, in which reduction of one type of risk (e.g., reduced SP risks by translocating fish among springs) may increase the other type of risk (i.e., increased OD risks due to mixing of divergent gene pools) (Figure 3.1). Fluker et al. (2010, p. 2276) recommend that two of the four populations be maintained in isolation (specific recommendations for the other two populations are not mentioned), implicitly favoring an OD-risk-averse

94

strategy over an SP-risk-averse strategy. However, although they acknowledge that populations face both risk types, they cite no explicit criteria for how these risks were ranked or explanation for why OD risks received primacy. Although Fluker et al.'s (2010) recommendations may be the best strategy for the species' conservation, they could be questioned for the lack of a scientifically-based rationale. The situation could become controversial if, for example, maintenance of isolation prevented the deployment of recovery options (e.g., translocation) that could lead to species down-listing.

In another example, Yamamoto et al. (2006) retrospectively considered the values of having translocated white-spotted char (*Salvelinus leucomaenis*) from downstream to upstream of erosion-control dams on two rivers in Hokkaido, Japan. Prior to translocation, above-dam populations exhibited low genetic diversity and very small N_es and were highly vulnerable to extinction (Yamamoto et al. 2006; Morita and Yamamoto 2002), thus exhibiting high SP risks. However, populations above and below dams exhibited differential migratory behaviors, growth rates, and ages at maturity (reviewed in Yamamoto et al. 2006), indicative of elevated OD risks. Following the translocation, Yamamoto et al. (2006) recorded dramatic increases in neutral marker diversity in above-dam populations, interpreting this as a successful genetic rescue and implicitly as a reduction in SP risks. However, as in the previous example, Yamamoto et al. (2006) did not describe whether or how SP and OD risks were weighed prior to undertaking the translocation. Moreover, although they caution that future translocations should consider the evolutionary relationships among Japanese *S. leucomaenis* populations, which in some cases may be distant (Yamamoto et al. 2004), they offer no suggestions for how such data should be used. As in the previous example, the management strategy employed by Yamamoto et al. (2006) may have been the best course of action, but no analysis was conducted with which to assess this possibility. In their study, above- and below-dam populations had been isolated for < 10 generations, which should correspond to low OD risk (Frankham et al. 2011). However, barriers in other Japanese rivers may have been in place much longer, and even in the absence of barriers, salmonids can develop significant life-history variation over small spatial extents (Neville et al. 2006). These factors could increase OD risks, indicating the danger of blindly applying Yamamoto et al.'s (2006) recommendations to other systems.

Although I found SP and OD risks to be inversely related for *P. rex* (see below), I do not expect this to be a general feature of risk for other species. A quantitative risk assessment may reveal that some individual populations of a threatened species are at low risk for both types, occupy the lower left quadrant of Figure 3.1, and therefore are in no immediate need of management intervention. In contrast, populations of species such as *E. nuchale* (see above) and other species that now occupy tiny, isolated remnants of historically much-larger ranges (e.g., Southeastern Fishes Council 2008) likely are at high risk both for SP and OD factors. Management strategies for this type of population, which occupies the top right quadrant of Figure 3.1, are difficult to prescribe. Conservation of such a population will involve careful scientific evaluation of whether it will persist without human intervention (e.g., Hedrick 1995), and if not, non-scientific value judgments about which evolutionary units (populations, evolutionarily significant units, or species) are most important to preserve (Angermeier 2001).

Although my risk assessment framework represents a step forward from previous, mostly implicit approaches, there are several ways that the framework could be improved. First, additional or alternative metrics could be used to assess risk level. Based on data availability for *P. rex*, I used metrics developed from geographic and molecular genetic data, using these as surrogates for direct measures of population size and adaptive divergence. However, if the latter data are available, their use would provide preferable measures of risk. Such data could be acquired from long-term field studies of demography and breeding structure, experimental studies of the fitness-divergence relationships in inter-crossed progeny, and studies of variation at fitness-related genes and quantitative traits (e.g., Vrijenhoek 1996; Fraser et al. 2010). Second, the criteria themselves could be refined, based on improved, taxon-specific models relating metric values to persistence and adaptive potential over time-frames of interest. Few such models exist for stream fishes, and those that do were developed for salmonids (e.g., Waples 1990; Emlen 1990; Hilderbrand and Kershner 2000; Rieman and Allendorf 2001). Third, future risk assessments could incorporate the uncertainty associated with metric values and criteria, or perhaps apply a weighting scheme to risk types based on the importance placed on those features. One way of incorporating this

96

uncertainty and weighting would be through the use of Bayesian belief networks (e.g., Peterson et al. 2008).

Designing a risk-averse conservation strategy for Percina rex

Like many threatened stream fishes, *Percina rex* persists in a small number of isolated remnant habitats. The demographic and evolutionary viability of these populations is at risk from processes inherent to small populations, including demographic and environmental stochasticity, inbreeding depression, and the loss of adaptive potential to drift. Government agencies tasked with recovering *P. rex* must decide which of a limited number of available recovery strategies to pursue to ensure and enhance the viability of these seven populations. Habitat restoration could potentially increase population vital rates and boost population size, but these benefits likely would accrue slowly, perhaps over timescales too long to prevent extirpation of the smallest populations. Removal of dispersal barriers between populations could provide demographic and genetic rescue, but would require the removal of major dams and reservoirs, which would be politically intractable and could potentially create new conservation problems (Stanley and Doyle 2003; Fausch et al. 2009). Upstream range expansion of populations is precluded by the species' preference for large streams. Augmentation of populations via translocations among localities could potentially increase population size and inject new genetic diversity, but could have unintended negative consequences such as OD. Tension between these two types of risks – SP risks due to maintenance of isolation versus OD risks due to restoration of connectivity – is the key challenge to conservation planning for *P. rex*. This tension permeates the management of many other threatened stream fishes as well (George et al. 2009).

Based on my analysis of SP and OD risks to *P. rex* populations, a set of straightforward, scientifically defensible management guidelines emerge. Critically, the relative risks due to maintained isolation and instituted translocation varied widely among populations, indicating that neither management option would be appropriate as a blanket policy. Only one population, occupying the lower section of the Smith River and its tributaries, appeared to be well insulated against both types of risk. This population tentatively requires no management intervention. The other six populations exhibited

97

moderate to high risk for one risk type or the other and may require management intervention. Fortunately, no population exhibited high risk for both risk types.

Populations GOOSE, OTTER, PIGG, and USMITH exhibit moderate to high SP risks due to small population sizes and geographic ranges restricted to a few stream channels. The population in Goose Creek is at particularly high risk, given its N_e of less than 50 individuals and restriction to a single stream. Both of these factors make GOOSE susceptible to extinction; the former because it increases the chance of inbreeding depression (Franklin 1980) and the latter because it increases the chance that a single catastrophic event kills all members of the population. Because none of these populations exhibit high OD risk, translocation is a reasonable option for increasing population viability. Each population is closely related genetically and ecologically to at least one other population that occurs in the same basin and physiography (Table 3.2). Managed gene flow between populations within these basin-physiography units may be prudent.

At the other end of the risk-space spectrum, the population occupying the Nottoway River exhibited moderate OD risk due to a lack of close relatives with which gene flow could be instituted. NOTT occurs in a unique basin and physiographic province (Coastal Plain) that is spatially distant from all other extant populations. Habitat characteristics of Coastal Plain streams, such as geology, stream geomorphology, and water chemistry, differ significantly from those of streams elsewhere in *P. rex*'s range (Rosenberger 2002). Furthermore, adult *P. rex* occupy different habitat configurations in the Coastal Plain than in other physiographies (Rosenberger 2002). Although I lacked quantitative criteria to incorporate these environmental and phenotypic differences into the risk-assessment framework, they provide further circumstantial evidence that NOTT could be ecologically divergent from other populations. It therefore may be prudent to maintain the isolation of this population. The population occupying the upper Roanoke River was on the threshold that discriminated low from moderate OD risks. Because this population occupies a unique physiographic province (Valley and Ridge), a tentative strategy of maintained isolation may be prudent. However, risks to UROAN of both types are relatively

low, so if other translocation sources within the Roanoke basin fail to accomplish restoration objectives, UROAN could be considered as an alternative translocation source for this basin.

As with any management action, the results of translocation or isolation strategies should be monitored over time to determine whether they meet conservation objectives. For isolation, the objective is to maintain population viability with no increase in SP risks. For translocation, the objective is to enhance population viability and reduce SP risks. The effective size of most populations did not meet the "50-500" criterion (Franklin 1980) believed necessary for long-term viability. Until a more species-specific criterion is developed, I suggest that this criterion be used as a minimum benchmark for assessing population recovery. Assessing progress toward this benchmark and others will require development of a rigorous monitoring program for *P. rex* populations. These data plus others collected should periodically be used to re-evaluate SP and OD risks for the species, ascertain whether they have changed, and revise management strategies as necessary.

Conclusions

Although conservation decisions cannot be distilled down to a simple algorithm, risk-assessment frameworks are useful for measuring and communicating the benefits and costs of alternative management options. A common decision faced by managers of small, fragmented populations is whether or not to intervene and attempt to modify population dynamics via augmentation (Jones et al. 2006; George et al. 2009). A transparent assessment of risk, based on best available scientific data, is the key to making such decisions. Better quantitative models relating SP and OD risk factors to persistence are needed for most species of stream fish. Incorporation of these relationships into frameworks like the one presented herein can help guide conservation choices.

ACKNOWLEDGMENTS

REFERENCES

Allendorf, F.W., and R.F. Leary. 1986. Heterozygosity and fitness in natural populations of animals. Pages 57-76 in M.E. Soule, editor. Conservation Biology: The Science of Scarcity and Diversity. Sinauer Associates Inc., Sunderland, MA.

Allendorf, F.W., and nine other authors. 1997. Prioritizing Pacific salmon stocks for conservation. Conservation Biology 11:140-152.

Angermeier, P.L. 2001. Reply to Povilitis: another conservation conundrum. Conservation Biology 15:536-538.

Angermeier, P.L., and M. R. Winston. 1999. Characterizing community diversity across Virginia landscapes: prerequisite for conservation. Ecological Applications 9:335-349.

Beaumont, M.A., W. Zhang, and D.J. Balding. 2002. Approximate Bayesian computation in population genetics. Genetics 162:2025-2035.

Berger, J. 1990. Persistence of different-sized populations: an empirical assessment of rapid extinctions in bighorn sheep. Conservation Biology 4:91-98.

Bolnick, D.I., and T.J. Near. 2005. Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1754-1767.

Caughley, G. 1994. Directions in conservation biology. Journal of Animal Ecology 63:215-244.

Coltman, D.W., and J. Slate. 2003. Microsatellite measures of inbreeding: a meta-analysis. Evolution 57:971-983.

Crandall, K.A., O.R.P. Binanda-Emonds, G.M. Mace, and R.K. Wayne. 2000. Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution 15:290-295.

Dunham, J.B., S.B. Adams, R.E. Schroeter, and D.C. Novinger. 2002. Alien invasions in aquatic ecosystems: toward an understanding of brook trout invasions and their potential impacts on inland cutthroat trout. Reviews in Fish Biology and Fisheries 12:373-339.

Edmands, S. 2002. Does parental divergence predict reproductive compatibility? Trends in Ecology and Evolution 17:520-527.

Edmands, S. 2007. Between a rock and hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Molecular Ecology 16:463-475.

Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics 5:435-445.

Emlen, J.M. 1991. Heterosis and outbreeding depression: a multilocus model and an application to salmon production. Fisheries Research 12:187-212.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin version 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Fagan, W.F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. Ecology 83: 3243-3249.

Fausch, K.D., B.E. Reiman, J.B. Dunham, M.K. Young, and D.P. Peterson. 2009. Invasion versus isolation: trade-offs in managing native salmonids with barriers to upstream movement. Conservation Biology 23:859-870.

Fluker, B.L., B.R. Kuhajda, N.J. Lang, and P.M. Harris. 2010. Low genetic diversity and small long-term population sizes in the spring endemic watercress darter, *Etheostoma nuchale*. Conservation Genetics 11:2267-2279.

Francis, R.I.C.C., and R. Shotton. 1997. "Risk" in fisheries management: a review. Canadian Journal of Fisheries and Aquatic Sciences 54:1699-1715.

Frankham, R., J.D. Ballou, and D.A. Briscoe. 2010. Introduction to Conservation Genetics, second edition. Cambridge University Press, Cambridge, U.K.

Frankham, R., and six other authors. 2011. Predicting the probability of outbreeding depression. Conservation Biology 25:465-475.

Franklin, I.R. 1980. Evolutionary change in small populations. Pages 135-149 in M.E. Soule, and B.A. Wilcox, editors. Conservation Biology: An Evolutionary-Ecological Perspective. Sinauer Associates Inc., Sunderland, MA.

Franklin, I.R., and R. Frankham. 1998. How large must populations be to retain evolutionary potential? Animal Conservation 1:69-73.

Fraser, D.J., A.L.S. Houde, P.V. Debes, P. O'Reilly, J.D. Eddington, and J.A. Hutchings. 2010. Consequences of farmed-wild hybridization across divergent wild populations and multiple traits in salmon. Ecological Applications 20:935-953.

George, A.L., B.R. Kuhajda, J.D. Williams, M.A. Cantrell, P.L. Rakes, and J.R. Shute. 2009. Guidelines for propagation and translocation for freshwater fish conservation. Fisheries 34:529-545.

Groom, M.J., G.K. Meffe, and C.R. Carroll. 2006. Principles of Conservation Biology, Third Edition. Sinauer Associates Inc., Sunderland, MA.

Hedrick, P.W. 1995. Gene flow and genetic restoration: the Florida panther as a case study. Conservation Biology 9:996-1007.

Hedrick, P.W. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution 53:313-318.

Hedrick, P.W. 2009. Genetics of Populations, 4th Edition. Jones and Bartlett Publishers, Sudbury, MA.

Hedrick, P.W., and R. Fredrickson. 2010. Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. Conservation Genetics 11:615-626.

Hilderbrand, R.H., and J.L. Kershner. 2000. Conserving inland cutthroat trout in small streams: how much stream is enough? North American Journal of Fisheries Management 20:513-520.

Hudson, R.R., M. Slatkin, and W.P. Maddison. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics 132:583-589.

Jenkins, R.E., and N.M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, MD.

Jones, J.W., E.M. Hallerman, and R.J. Neves. 2006. Genetic management guidelines for captive propagation of freshwater mussels (Unionoidea). Journal of Shellfish Research 25:527-535.

Lahey, A.M., and P.L. Angermeier. 2007. Range-wide assessment of habitat suitability for Roanoke logperch (*Percina rex*). Final Report to Virginia Transportation Research Council, Charlottesville Virginia.

Librado, P. and J. Rozas. 2009. DNAsp version 5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

Lindley, S.T., and 11 other authors. 2007. Framework for assessing viability of threatened and endangered Chinook salmon and steelhead in the Sacramento-San Joaquin Basin. San Francisco Estuary and Watershed Science. Available online at http://repositories/cdlib.org/jmie/sfews/vol5/iss1/art4.

Lopes, J.S., and S. Boessenkool. 2010. The use of approximate Bayesian computation in conservation genetics and its application in a case study on yellow-eyed penguins. Conservation Genetics 11:421-433.

Lowe, W.H., and F.W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051.

Luikart, G., N. Ryman, D.A. Tallmon, M.K. Schwartz, and F.W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conservation Genetics 11:355-373.

McCauley, D.E., and M.J. Wade. 1981. The populational effects of inbreeding in *Tribolium*. Heredity 46:59-67.

McClelland, E.K., and K.A. Naish. 2007. What is the fitness outcome of crossing unrelated fish populations? A meta-analysis and an evaluation of future research directions. Conservation Genetics 8:397-416.

McElhany, P., M.H. Ruckelshaus, M.J. Ford, T.C. Wainwright, and E.P. Bjorkstedt. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. United States Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-42.

Mills, L. S., and F. W. Allendorf. 1996. The one-migrant-per-generation rule in conservation and management. Conservation Biology 10:1509-1518.

Minckley, W.L., P.C. Marsh, J.E. Deacon, T.E. Dowling, P.W. Hedrick, W.J. Matthews, and G. Mueller. 2003. A conservation plan for native fishes of the lower Colorado River. BioScience 53:219-232.

Morita, K., and S. Yamamoto. 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. Conservation Biology 16:1318-1323.

Nei, M. and A. Chakravarti. 1977. Drift variances of F_{ST} and G_{ST} statistics obtained from a finite number of isolated populations. Theoretical Population Biology 11:307-325.

Neville, H.M., J.B. Dunham, and M.M. Peacock. 2006. Landscape attributes and life history variability shape genetic structure of trout populations in a stream network. Landscape Ecology 21:901-916.

Peterson, D.P., B.E. Rieman, J.B. Dunham, K.D. Fausch, and M.K. Young. 2008. Analysis of trade-offs between threats of invasion by nonnative brook trout (*Salvelinus fontinalis*) and intentional isolation for native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*). Canadian Journal of Fisheries and Aquatic Sciences 65:557-573.

Pimm, S.L., G.J. Russell, J.L. Gittleman, and T.M. Brooks. 1995. The future of biodiversity. Science 269:347-350.

Pritchard, V.L., K. Jones, and D.E. Cowley. 2007. Genetic diversity within fragmented cutthroat trout populations. Transactions of the American Fisheries Society 136:606-623.

Reed, D.H., J.J. O'Grady, B.W. Brook, J.D. Ballou, and R. Frankham. 2003. Estimates of minimum viable population sizes for vertebrates and factors influencing those estimates. Biological Conservation 113:23-34.

Ricciardi, A. and J.B. Rasmussen. 1999. Extinction rates of North American freshwater fauna. Conservation Biology 13:1220-1222.

Rieman, B.E., and F.W. Allendorf. 2001. Effective population size and genetic conservation criteria for bull trout. North American Journal of Fisheries Management 21:756–764.

Roberts, J.H., and A.E. Rosenberger. 2008. Threatened fishes of the world: *Percina rex* (Jordan and Evermann 1889) (Percidae). Environmental Biology of Fishes 83:439-440.

Roberts, J.H., Angermeier, P.L., and E.M. Hallerman. 2009. Analysis of population genetics of Roanoke logperch (*Percina rex*) based on microsatellite markers. Final Report to U.S. Fish and Wildlife Service, Gloucester, VA.

Roberts, J.H., Angermeier, P.L., and E.M. Hallerman. 2010. Survey of mitochondrial DNA variation in endangered Roanoke logperch (*Percina rex*). Final Report to Virginia Department of Game and Inland Fisheries, Richmond, VA.

Rosenberger, A.E. 2007. An update to the Roanoke logperch Recovery Plan. Final Report to the U.S. Fish and Wildlife Service, Gloucester, VA.

Russell, S.T. 2003. Evolution of intrinsic post-zygotic reproductive isolation in fish. Annals of Zoology Fennici 40:321-329.

Skalski, G.T., J.B. Landis, M.J. Grose, and S.P. Hudman. 2008. Genetic structure of creek chub, a headwater minnow, in an impounded river system. Transactions of the American Fisheries Society 137:962-975.

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

Smith, K.L., and M.L. Jones. 2005. Watershed-level sampling effort requirements for determining riverine fish species composition. Canadian Journal of Fisheries and Aquatic Sciences 62:1580-1588.

Southeastern Fishes Council. 2008. The desperate dozen: fishes on the brink. Available online at: http://ichthyology.usm.edu/sfc/DesDoz_report.pdf.

Stanley E.H., and M.W. Doyle. 2003. Trading off: the ecological effects of dam removal. Frontiers in Ecology and the Environment 1:15-22.

Tallmon, D.A., G. Luikart, and R.S. Waples. 2004. The alluring simplicity and complex reality of genetic rescue. Trends in Ecology and Evolution 19:489-496.

Taylor, C.M., and M.L. Warren. 2001. Dynamics in species composition of stream fish assemblages: environmental variability and nested subsets. Ecology 82: 2320-2330.

Thompson, W.L., G.C. White, and C. Gowan. 1998. Monitoring Vertebrate Populations. Academic Press, San Diego, CA.

Townsend, C.R., M.R. Scarsbrook, and S. Doledec. 1997. The intermediate disturbance hypothesis, refugia, and biodiversity in streams. Limnology and Oceanography 42: 938-949.

United States Fish and Wildlife Service. 2000. USFWS-NMFS policy regarding controlled propagation of species listed under the ESA. U.S. Federal Register 65:56916-56922.

Vrijenhoek, R.C. 1996. Conservation genetics of North America desert fishes. Pages 367-397 in J.C. Avise and J.L. Hamrick, editors. Conservation Genetics: Case Histories from Nature. Chapman and Hall, New York, NY.

Waples, R.S. 1990. Conservation genetics of Pacific salmon. II: effective population size and the rate of loss of genetic variability. Journal of Heredity 81:267-276.

Weir, B.S. and C.C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.

Winston, M.R., C.M. Taylor, and J. Pigg. 1991. Upstream extirpation of four minnow species due to damming of a prairie stream. Transactions of the American Fisheries Society 120:98-105.

Xu, H., and Y-X. Fu. 2004. Estimating effective population size or mutation rate with microsatellites. Genetics 166:555-563.

Yamamoto, S., K. Morita, S. Kitano, K. Watanabe, I. Koizumi, K. Maekawa, and K. Takamura. 2004. Phylogeography of white-spotted charr (*Salvelinus leucomaenis*) inferred from mitochondrial DNA sequences. Zoological Society 21:229-240.

Yamamoto, S., K. Maekawa, T. Tamate, I. Koizumi, K. Hasegawa, and H. Kubota. 2006. Genetic
evaluation of translocation in artificially isolated populations of white-spotted charr (*Salvelinus leucomaenis*). Fisheries Research 78:352-358.

Young, M.K., and A.L. Harig. 2001. A critique of the recovery of greenback cutthroat trout.
Conservation Biology 15:1575-1584.

Table 3.1. Potential metrics by which levels of small-population (SP) and outbreeding-depression (OD) risk to stream-fish populations might be quantified. Only those metrics shown in bold typeface were used to assess risk to *Percina rex*. The depicted scoring criteria were used to assess risk to *P. rex*, but could be modified as needed for other species.

Risk type	Metric	Level of risk (score)		
		Low (1)	Moderate (3)	High (5)
SP	**Adult population size (*N*)**	> 7000	500-7000	< 500
	Effective population size (*N_e*)	> 500	50-500	< 50
	Inbreeding coefficient (*F*)	< 0.1	0.1-0.25	> 0.25
	Number of occupied stream channels	> 3	2-3	1
	Extinction risk from PVA	Low	Moderate	High
	Environmental stochasticity	Low	Moderate	High
	Population growth rate	Stable to increasing	Slowly decreasing	Rapidly decreasing
OD	**Sequence divergence (%)**	< 0.2	0.2-7	> 7
	Fluvial distance between populations (km)	< 80	80-160	> 160
	Time since isolation in…			
	…same environment	≤ 500 years		> 500 years
	…different environment	≤ 20 generations	> 20 generations	> 500 years
	Biogeographic overlap	Basin and physiography	Basin or physiography	Neither
	Phenotypic characters	Complete overlap	Partial overlap	No overlap
	Adaptive marker allele-frequencies	Complete overlap	Partial overlap	No overlap

Table 3.2. Estimated genetic characteristics of seven populations of *Percina rex*, based on analysis of the *ND2* mtDNA gene and 11 microsatellite loci. Entries include the sample size of individuals (n) and the estimated number of *ND2* haplotypes (K), segregating sites (S), haplotype diversity (H_d), and nucleotide diversity (π). For microsatellites, unbiased gene diversity (H_E) and allele richness (A) are estimated by averaging across loci. Estimates of female (N_{ef}) and total (N_e) effective population size are based on approximate Bayesian computation models (see text).

Population	n	K	Private haplotypes	S	H_d	π	N_{ef}	H_E	A	N_e
					ND2				Microsatellites	
UROAN	30	11	8	16	0.885	0.0031	280	0.83	10.1	1198
PIGG	30	2	0	4	0.460	0.0018	72	0.65	5.4	601
GOOSE	30	2	0	4	0.331	0.0013	20	0.53	3.8	11
OTTER	30	1	0	0	0.000	0.0000	5	0.56	3.0	99
LSMITH	30	2	1	4	0.515	0.0020	51	0.68	5.3	698
USMITH	30	3	0	5	0.393	0.0016	38	0.79	8.2	196
NOTT	30	2	2	1	0.497	0.0005	11	0.71	6.2	300
Total	210	15		22	0.775	0.0032				

109

Table 3.3. Estimates of genetic, temporal, and spatial distance between pairs of *Percina rex* populations. Explanations of genetic statistics are given in the text. Divergence time (t, in generations) is estimated from an isolation model, based on the F_{ST} and harmonic mean effective population size between populations estimated from microsatellite data.

		ND2		Microsatellites			Fluvial
		F_{ST}	% divergence	F_{ST}	R_{ST}	t	distance (km)
UROAN	PIGG	0.400	0.2	0.128	0.131	220	127.5
UROAN	GOOSE	0.512	0.3	0.204	0.215	10	111.5
UROAN	OTTER	0.697	0.4	0.176	0.256	71	132.1
UROAN	LSMITH	0.300	0.1	0.129	0.240	243	373.0
UROAN	USMITH	-0.002	0.0	0.043	0.164	30	469.7
UROAN	NOTT	0.586	0.3	0.148	0.170	154	824.3
PIGG	GOOSE	0.011	0.0	0.100	0.079	5	107.6
PIGG	OTTER	0.310	0.0	0.109	0.111	39	128.2
PIGG	LSMITH	0.269	0.1	0.263	0.202	394	369.1
PIGG	USMITH	0.486	0.2	0.159	0.118	103	465.8
PIGG	NOTT	0.753	0.4	0.261	0.238	243	820.4
GOOSE	OTTER	0.172	0.0	0.208	0.136	9	51.2
GOOSE	LSMITH	0.390	0.1	0.324	0.227	17	292.1
GOOSE	USMITH	0.617	0.3	0.212	0.105	10	388.8
GOOSE	NOTT	0.818	0.5	0.334	0.220	17	743.4
OTTER	LSMITH	0.649	0.2	0.316	0.311	132	278.9
OTTER	USMITH	0.823	0.4	0.199	0.186	59	375.6
OTTER	NOTT	0.954	0.6	0.327	0.332	118	730.2
LSMITH	USMITH	0.364	0.1	0.105	0.067	68	30.4
LSMITH	NOTT	0.714	0.4	0.201	0.284	188	725.7
USMITH	NOTT	0.719	0.3	0.160	0.110	83	822.4

Table 3.4. Estimated values and resultant risk scores for metrics used to assess relative small-population (SP) and outbreeding-depression (OD) risks to seven populations of *Percina rex*. Each metric was scored on a three-category scale (1 = low risk, 3 = moderate risk, 5 = high risk) and then scores were averaged across metrics within risk types. Physiographies include Valley and Ridge (VR), Piedmont (PD), and Coastal Plain (CP).

Metric	Metric value (risk score)						
	UROAN	PIGG	GOOSE	OTTER	LSMITH	USMITH	NOTT
Adult population size	13884 (1)	6860 (3)	1616 (3)	1586 (3)	9362 (1)	2497 (3)	16686 (1)
Mean effective population size	879 (1)	373 (3)	26 (5)	55 (3)	400 (3)	136 (3)	161 (3)
Increase of F over next 100 years	0.01 (1)	0.03 (1)	0.35 (5)	0.18 (3)	0.03 (1)	0.08 (1)	0.07 (1)
Number of occupied stream channels	4 (1)	2 (3)	1 (5)	2 (3)	6 (1)	2 (3)	5 (1)
Overall SP risk score	1.0	2.5	4.5	3.0	1.5	2.5	1.5
Minimum sequence divergence (%)	0.0 (1)	0.0 (1)	0.0 (1)	0.0 (1)	0.1 (1)	0.0 (1)	0.3 (3)
Distance to closest population (km)	111.5 (3)	107.6 (3)	51.2 (1)	51.2 (1)	30.4 (1)	30.4 (1)	725.7 (5)
Minimum divergence time (gen/yr)	10/45 (1)	5/20 (1)	5/20 (1)	9/42 (1)	17/76 (1)	10/45 (1)	17/78 (1)
Occupied basin/physiography	Roanoke/VR (3)	Roanoke/PD (1)	Roanoke/PD (1)	Roanoke/PD (1)	Dan/PD (1)	Dan/PD (1)	Nottoway/CP (5)
Overall OD risk score	2.0	1.5	1.0	1.0	1.0	1.0	3.5
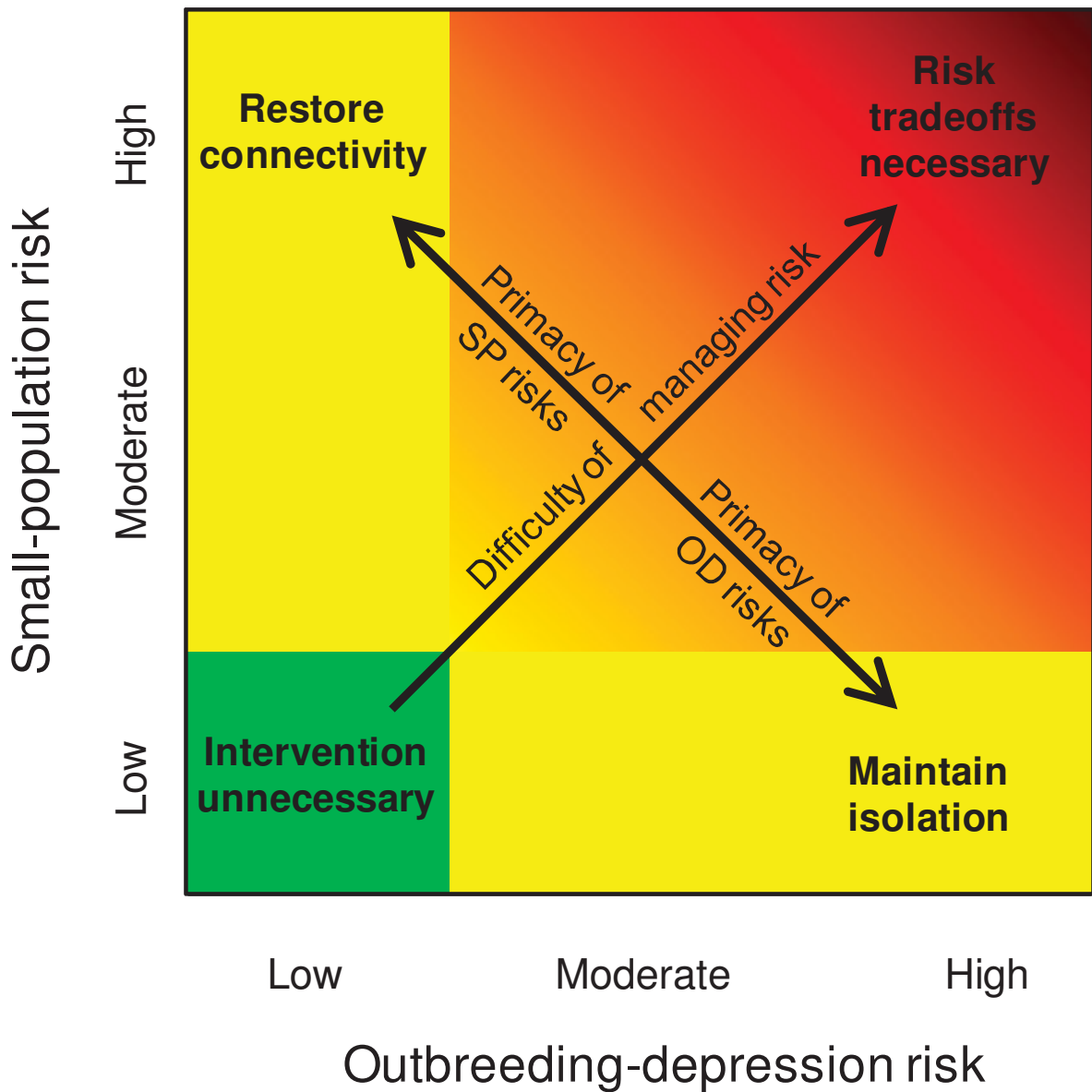
Figure 3.1 Conceptual model of management prescriptions emanating from an assessment of small-population (SP) and outbreeding-depression (OD) risks to fragmented populations. When populations have low risk of both risk types (green area), management intervention is unnecessary. When populations have moderate to high risk for only one risk type (yellow areas), the optimal management strategy (restore connectivity versus maintain isolation) is straightforward. However, when populations are at high risk for both risk types (red area), the optimal management strategy is unclear and likely to involve tradeoffs of one form of risk for another.
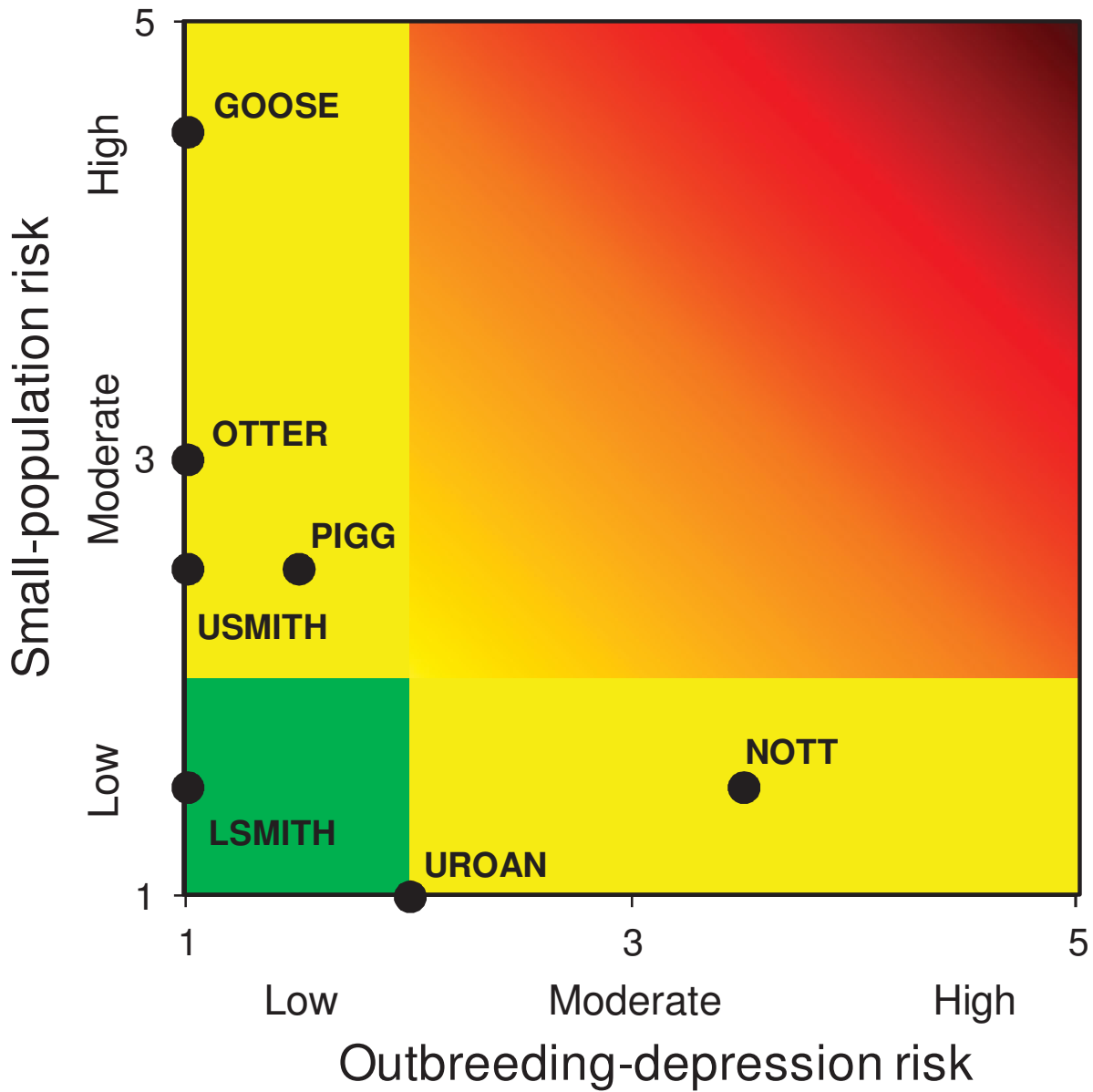
Figure 3.2. Results of an assessment of small-population and outbreeding-depression risks to seven populations of *Percina rex* (filled circles). Population coordinates are based on the mean risk score across four metrics within each risk type (see Table 3.4). Overall risk ranged from 1 (low risk across all metrics) to 5 (high risk across all metrics). The meaning of background colors is explained in Figure 3.1.

CHAPTER 4: Influences of urbanization, stream size, and fragmentation on the genetic diversity and differentiation of two species of stream fish

ABSTRACT

Human alterations of the landscape can profoundly affect the structure and function of rivers and streams. Population genetic data are an underutilized tool for testing hypotheses about the combined influences of land use, stream size, and habitat fragmentation on population dynamics of stream biota. I examined spatial patterns of variation in microsatellite DNA diversity and differentiation within redline (*Etheostoma rufilineatum*) and greenside darters (*E. blennioides*), two common species of stream fish of the upper Tennessee River basin, a region characterized by extensive hydrologic alteration and land-use change. Random Forest multiple regression models indicated that genetic diversity of both species was positively related to stream size and negatively related to hydrologic isolation by reservoirs, consistent with island biogeographic expectations that smaller, more isolated patches harbor smaller and/or more variable populations. Both species also exhibited pronounced genetic differentiation among sites fragmented by dams and reservoirs, which largely overrode any signal of genetic isolation-by-distance. Indices of urban and agricultural land use were calculated for watersheds surrounding darter sample sites. Influences of these land-use predictor variables varied somewhat among genetic response variables and between species. However, indices of urbanization (% urban land use, % impervious surfaces, and road density) were consistently negatively related to genetic indices of darter population size and stability. My results demonstrate that urbanization can negatively impact the demography of common fish species, and that genetic monitoring can provide information on population status that complements information provided by conventional biological monitoring.

INTRODUCTION

Human alterations of the landscape can profoundly affect the structure and function of rivers and streams (Allan 2004). Conversion from forested to agricultural or urban land use in a watershed can result in increased fine-sediment and nutrient loading, destabilized stream channels, increased chemical pollution, reduced habitat complexity, increased water temperatures, and greater streamflow extremes

114

(Lenat 1984; Osborne and Wiley 1988; Walsh et al. 2005). In addition, direct modifications of hydrology, such as the impoundment or channelization of free-flowing rivers, dramatically alter the flow of nutrients, sediments, food, and habitat-forming energy in these systems (Vannote et al. 1980; Junk et al. 1989; Poff et al. 1997).

Stream organisms are highly responsive to these stressors, and provide useful indicators of ecosystem change. Unlike snapshot measures of water or habitat quality, which may miss rare pollution events or interactions among multiple stressors, biotic measures provide a more synthetic view of conditions over the temporal scales of organismal lifespans. Effects of stressors on biota can be measured at any level of organization, including individual physiology, fitness, and survival (e.g., Sullivan and Lydy 1999; Sugg et al. 1995; Larno et al. 2001), population size, age structure, and genetic diversity (Heithaus and Laushman 1997; Peoples and Frimpong, in press), and community taxonomic and functional composition (Hilsenhoff 1998; Karr et al. 1986). However, most previous biomonitoring has focused either at the individual or community level, rather than on populations. Individual-based studies are tractable in laboratory or mesocosm experiments, and provide a direct link between particular pollutants and fitness consequences. However, individual-based studies have limited ability to detect the impacts of multiple, diffuse stressors that enter streams across large areas. Community-based approaches are more useful in this regard, because they are relatively easy to conduct across large areas and because certain taxa and guilds respond predictably to diffuse stressors such as sedimentation, food-web rearrangement, and habitat loss (Vannote et al. 1980; Berkman and Rabeni 1987; Townsend and Hildrew 1994).

Despite its popularity, the use of biological community data to characterize stream condition is limited in its utility for drawing strong inferences about how human activities impact stream biota. Species-composition data provide only an indirect proxy for the population dynamics that resulted in particular species occupying or not occupying a location. Knowing the demographic histories of individual populations, and how they relate to landscape alteration, would provide a more mechanistic understanding of anthropogenic impacts. A species' absence from a location may be due to various

115

factors unrelated to present landscape conditions, such as biogeographic history (Sheldon 1988), dispersal limitations that prevent colonization, natural lack of suitable habitat (e.g., due to species' thermal or stream-size preferences), and legacy effects from historical land use conditions (Harding et al. 1998). Such confounding factors make it difficult to calibrate community expectations for any given location and complicate comparisons among locations. Finally, extensive migration among localities can cloud the relationship between local community structure and local habitat conditions, and the spatial domain over which community samples convey useful information varies because mobility and migratory habitats vary widely among taxa (Bunn and Hughes 1997; Hitt and Angermeier 2008).

Examination of the population genetic characteristics of biota provides a novel, instructive complement to conventional bioassessments. Genetic data potentially provide two categories of information about biotic responses to anthropogenic stressors: 1) evolutionary responses to novel selective or mutagenic forces that drive allele-frequency changes in functional genes, and 2) demographic responses to changes in effective population size and migration rate that drive allele-frequency changes in selectively neutral genes (van Straalen and Timmermans 2002). The first category of responses is well explored with regard to the expression of stressor-resistant genes in aquatic populations exposed to severe stressors (e.g., toxins, temperature extremes) (Sullivan and Lydy 1999; Vrijenhoek et al. 1992). The second category of responses has been understudied, possibly because selectively neutral, rapidly mutating genetic markers such as microsatellites have only recently become widely available. Such markers enable the testing of spatially and temporally finer-scaled hypotheses than was possible using older markers such as allozymes. For example, statistics based on the distribution and diversity of microsatellite alleles within and among populations can indicate population size, population stability, and dispersal rate (Schwartz et al. 2006; Lowe and Allendorf 2010; Luikart et al. 2010). Such parameters are laborious to estimate using traditional field methods (Thompson et al. 1998), but are key elements of population persistence.

Genetic methods also exhibit several advantages over community-based approaches. First, genetic approaches examine intra-specific variation, which controls for potential inter-specific variation in

116

stressor tolerance, biogeographic history, dispersal limitations, or habitat preferences. Second, because genetic variation can reveal signatures of population decline (e.g., Garza and Williamson 2001), genetic methods may detect impacts earlier than community-based methods that are based on snapshot measures of abundance or presence/absence (i.e., prior to population extirpation), if the focal species are relatively sensitive to impact. Third, like community measures, genetic measures are potentially affected by off-site conditions. However, with genetic data, this zone of influence can be quantified directly by delineating population genetic structure and spatial extent. Collections made at two locations inferred to be part of the same population are not demographically independent, and therefore will not convey statistically independent information about biotic condition.

Like other biological indicators, population genetic measures must be calibrated based on landscape features extrinsic to land use (Osborne and Wiley 1992; Hitt and Angermeier 2008). Contemporary aquatic habitats often are fragmented into patches of varying size, much like the oceanic islands of Island Biogeography Theory (MacArthur and Wilson 1967). The size and isolation of these habitat fragments strongly influences the expected size, stability, connectivity, hence the expected levels of genetic diversity and differentiation, of constituent populations (Winston et al. 1991; Hilderbrand and Kershner 2000; Jones et al. 2000; Morita and Yamamoto 2002). For example, small streams generally exhibit greater environmental variability over time than large streams (Dunn and Leopold 1978; Horwitz 1978; Poff and Ward 1989). All else being equal, populations occupying small streams therefore should be smaller and more variable than populations occupying larger streams (Gotelli and Taylor 1999; Taylor and Warren 2001). Similarly, populations in isolated areas should exhibit lower immigration and demographic rescue than populations in connected areas. For fishes, connectivity among streams generally is high within unimpounded watersheds (Gorman 1986; Bessert and Ortí 2008; see Chapter 1), but low near physical barriers such as dams and waterfalls (Hänfling and Weetman 2006; Neville et al. 2006; Pritchard et al. 2007; Raeymaekers et al. 2008; Whiteley et al. 2010). Even unsuitable lentic habitats can function as barriers, elevating genetic differentiation and lowering genetic diversity (Mitchell et al. 2002; Skalski et al. 2008). Given that patch size and isolation influence the genetic diversity of

populations, these effects must be accounted for when testing hypotheses about land-use effects on stream biota.

In this study, I tested the influences, after accounting for stream size and hydrologic isolation, of a suite of land-use variables on genetic diversity and differentiation in populations of two common species of stream fish occupying the upper Tennessee River basin (UTRB). The UTRB is a biologically distinctive region of the eastern United States encompassing over 55,000 km^2 and portions of three major physiographic provinces (Ridge and Valley, Blue Ridge, and Appalachian Plateau) (Hampson et al. 2000; Figure 1). The UTRB encompasses many freshwater habitats, including creeks, rivers, springs, and impoundments (Etnier and Starnes 1993). Concomitant with this habitat diversity, the UTRB supports a globally outstanding diversity of freshwater species, many of which are threatened (Abell et al. 2000). Among the most serious threats are the loss and fragmentation of free-flowing habitats due to hydrologic alteration and impoundment, as well as reductions in habitat quality due to agricultural and urban development (Neves and Angermeier 1990; Mattson and Angermeier 2007). Large impoundments in the UTRB were completed between 1911 and 1979 (most between the 1930s and 1950s; Etnier and Starnes 1994). Agricultural development peaked in the early 20th century and has been decreasing since then, whereas urbanization generally has been increasing since the middle of the 20th century (Wear and Bolstad 1998; Diamond et al. 2002). The UTRB makes a good choice for my analysis, because of its widespread fragmentation, agricultural, and urban impacts, and its interest to managers charged with developing conservation plans for endangered species (Neves and Angermeier 1990).

My two study species were selected because of their potential for sensitivity to land-use variation and their broad distribution within the UTRB, which made it possible to set up a range of land use, stream size, and fragmentation contrasts. Redline (*Etheostoma rufilineatum*) and greenside (*E. blennioides*) darters are benthic, small-bodied (< 166-mm-long) members of the speciose North American darter subfamily (Percidae: Etheostomatinae). Darters are disproportionately imperiled by habitat loss and fragmentation (Etnier 1997; Jelks et al. 2008), presumably because most require clean gravel substrate for spawning and feeding, which renders them intolerant of silt deposition. *E. rufilineatum* is endemic to the

118

Tennessee and Cumberland river systems, whereas *E. blennioides* occurs widely throughout the Ohio, southeastern Great Lakes, Missouri, and Ouachita river systems (Etnier and Starnes 1993). Within the UTRB, both species occur throughout the Ridge and Valley province in large creeks and medium-sized rivers. Adults are most common in swift, rocky riffles and flowing pools, whereas juveniles prefer slower pools and stream margins. Although *E. blennioides* resides in glacial lakes in northern portions of its range, neither species occupies lentic habitats in UTRB impoundments. Both species mature in one to two years, live for four to five years, and spawn in the spring in sandy portions of riffles (Etnier and Starnes 1993; Jenkins and Burkhead 1994).

I hypothesized that genetic diversity in both species would be higher, and differentiation lower, in larger than smaller streams, because larger waterbodies would harbor larger, more stable and more connected populations. I also hypothesized that both species would exhibit genetic signals of fragmentation by dams and reservoirs. I predicted that urban and agricultural land use would correlate negatively with genetic diversity and positively with genetic differentiation, but I had no *a priori* expectations about which type of land use would best predict population genetic patterns. However, I hypothesized that, due to their ecological similarity, both species would respond similarly to a given landscape variable.

METHODS

Fish sampling

I collected fin clips non-lethally from darters at 23 sites distributed throughout the UTRB (Table 4.1; Figure 4.1). Each site was 100-200 m long and located on riffle-run habitat. At 22 of these sites, 23-24 adult *E. rufilineatum* individuals were sampled. Because they were more difficult to collect, fewer *E. blennioides* (i.e., 12-25 individuals) were sampled from 10 of the 23 sites (Table 4.1). Fish were captured with a Smith-Root direct-current backpack electrofisher and seines or dipnets, fin-clipped, and returned alive to the stream. Tissue samples were dried in paper envelopes and stored at -20 ℃ until DNA extraction.

DNA isolation and genotyping

I extracted template DNA from tissue samples using a Pure Gene DNA Extraction Core Kit A (Gentra Systems, Minneapolis, Minnesota, USA). At the time of the study, no microsatellite loci had been identified in either species, so I tested existing microsatellite DNA primers that had been developed for other darter species and utilized those that exhibited amplification success. Candidate markers included all loci listed in DeWoody et al. (2000; *E. olmstedi*), Porter et al. (2002; *E. virgatum*), and Gabel et al. (2008; *E. scotti*). Based on preliminary screening of a subset of individuals, the markers that demonstrated consistent amplification success and lack of null alleles for *E. rufilineatum* were *Cv12*, *Eo6*, *Esc18*, *Esc26*, and *Esc132*, whereas the markers suitable for *E. blennioides* were *Cv09*, *Cv12*, *Cv24*, *Eo6*, *Eo9*, *Esc26*, and *Esc132*. All sampled individuals were genotyped at these loci. Forward primers were fluorescently labeled (Applied Biosystems, Inc., Foster City, California, USA), and PCR was conducted in a MyCycler Thermal Cycler (BioRad, Hercules, California, USA; Appendix H). Amplification products were separated in an ABI 3130 automated sequencer and sized using GENEMAPPER 3.5 and a LIZ or ROX size standard (Applied Biosystems, Inc., Foster City, California, USA).

Estimation of genetic statistics

I tested for Hardy-Weinberg and linkage equilibrium for each species at each site in the program ARLEQUIN 3.11 (Excoffier et al. 2005). Hardy-Weinberg tests employed 10^5 recorded Markov-Chain-Monte-Carlo (MCMC) chains following a burn-in of 10^3 chains, whereas linkage tests employed 10^4 randomizations. To reduce the risk of type-I error, test results were evaluated at $\alpha = 0.01$.

As relative indices of population size and stability, I used FSTAT 2.9.3.2 (Goudet 2002) to estimate, for each site, the mean across loci of Nei's unbiased gene diversity (H_E) and allele richness accounting for sample size (A). I used ARLEQUIN to estimate the mean across loci of M, the ratio of the number of alleles to the size-range of alleles at a locus. The M ratio decreases after a population bottleneck (Garza and Williamson 2001), thereby providing a relative index of population stability. I quantified genetic differentiation by estimating F_{ST} values (Weir and Cockerham 1984) between pairs of

120

sites in ARLEQUIN. For each site-pair, I also tested the null hypothesis that $F_{ST} = 0$ via 10^4 permutations of individuals among populations, evaluating test results at $\alpha = 0.05$.

Landscape genetic analyses

I used multiple regression models to test the relative influences of hypothesized landscape characteristics on the genetic diversity and differentiation of *E. rufilineatum* and *E. blennioides* populations. In the first set of models, developed to explain variation in H_E, A, and M, each site was treated as an observation. However, sites belonging to the same population (i.e., with $F_{ST} \approx 0$) were presumed to be non-independent, so for diversity models I retained only the most-downstream site from each population. Candidate regressors measured stream size, hydrologic isolation, or watershed land-use of sites (Table 4.1, Appendix I). To index stream size, I used the U.S. Geological Survey (USGS) National Hydrography Dataset to estimate the watershed area upstream of each site. Isolation was indexed by measuring the distance along the stream from a site to the nearest downstream reservoir in Google Earth 5.1 (http://www.google.com/earth).

Ten regressor variables, characterizing land use near sites, were developed based on information from a recent USGS Aquatic Gap conservation assessment of the UTRB (Angermeier et al. 2009; see Appendix I). I first determined the USGS 12-digit hydrologic unit (HU12) in which each site was located, and then for selected HU12s, I summarized the prevalence of land uses deemed threatening to aquatic populations (Mattson and Angermeier 2007). Agricultural, developed, and forested area-based data were obtained from the 2001 National Land Cover Database (NLCD), summarized from 30-m^2 cells, and converted to percentages of total HU12 area. Of agricultural categories, I retained as separate regressors "cultivated crops" and "pasture/hay". The areas contained in "low-", "medium-", and "high-intensity developed" categories were summed as a single urban development variable. Likewise, a single forested-area variable summed percentages across multiple forested categories ("deciduous", "evergreen", and "mixed") and the "shrub/scrub" category. Data on percent impervious surfaces at the 30-m^2 resolution were obtained from the NLCD set. From the U.S. Census Bureau, I obtained data on road density and human population size within each HU12. I used the year 2000 census to estimate current

population size and the proportional change between 1960 and 2000 to estimate population growth rate. The number of active National Pollutant Discharge Elimination System (NPDES) permit sites (e.g., for industrial, agricultural, or sewage discharges) was obtained from the U.S. Environmental Protection Agency (http://www.epa.gov/enviro/index.html). Finally, I estimated the percentage of each HU12 located in protected conservation areas using data from the Aquatic GAP assessment; all GAP-category-1, -2, and -3 lands were considered protected (Angermeier et al. 2009).

A second set of models, developed to explain variation in F_{ST}, treated each pair-wise site comparison as an observation (n = 231 and 45 for *E. rufilineatum* and *E. blennioides*, respectively). Candidate regressors measured stream size of sites or hydrologic isolation between sites. Stream size was indexed by calculating the arithmetic mean upstream watershed area of each site pair. Isolation was measured in two ways. First, I estimated the distance between sites along stream channels in Google Earth. Second, I developed metrics to account for variation in the presumed permeability of stream reaches to movement by darters. Connectivity varied between pairs of sites, in that sites could be connected by unimpeded riverine connections, separated by the upper reaches of an impoundment (i.e., if located in adjacent tributaries that flow into the same reservoir), or separated by a major dam. Three alternative ordinal regressors, each consisting of different combinations of these connectivity scenarios, were developed for testing in *E. rufilineatum* models: a) solely riverine connection (assigned "0") *versus* reservoir or dam present (assigned "1"), b) dam absent (assigned "0") *versus* dam present (assigned "1"), and c) solely riverine connection (assigned "0") *versus* reservoir but no dam present (assigned "1") *versus* dam present (assigned "2"). In the case of *E. blennioides*, only one solely riverine connection was sampled, so only the second of these scales was tested.

Regression models were built using the Random Forest (RF) approach, an extension of classification and regression trees (Breiman 2001). Unlike classical regression techniques, which seek to explain a response variable using linear combinations of predictor variables, tree-based methods build predictive models by recursively partitioning the data into successively smaller groups based on a series of binary splitting rules defined by individual predictor variables (De'ath and Fabricius 2000). In the case

of regression trees, splitting rules are designed to minimize the within-group sum-of-squares of a

continuous response variable. Tree-based methods facilitate the modeling of correlated predictor

variables and non-linear, non-additive, and/or hierarchical relationships, which may be common in

landscape ecological studies (Allan 2004; Prasad et al. 2006). Whereas a regression tree model attempts

to find a single best predictive tree based on the entire dataset, an RF fits many trees to bootstrap samples

of the dataset and then combines predictions across all of the trees, a convention that reduces

classification bias and model over-fitting (Cutler et al. 2007).

RF models were built to explain variation in the three genetic diversity statistics (H_E, A, and M)

and the differentiation statistic F_{ST} for each of the two species (eight models total). Models were built

using the *randomForest* statistical package in *R* 2.10.0 (*R* Development Core Team 2004). Each RF

model employed 10^4 bootstrapped trees. The only other input parameter in RF is the number of predictor

variables to be evaluated at each tree split (*mtry*), and no firm criteria exist for selecting this value (Prasad

et al. 2006). Therefore, for each model, I evaluated *mtry* values of one to five and utilized the value that

maximized the percentage of explained variance. At any rate, the use of different *mtry* values had no

substantive effect on the identity or rank order of regressors deemed important. Regressor importance

was measured by the percentage increase in mean squared prediction error when a regressor's values were

randomly permuted among observations. Thus, a higher value indicates greater contribution of a

regressor variable to the model. Regressors with importance values > 10% were further interpreted using

partial dependence plots, which illustrate the effect of a regressor on the mean of the response variable

when all other regressors are averaged out (Friedman 2001).

RESULTS

Hardy-Weinberg equilibrium was rejected in only 1 of 110 tests for *E. rufilineatum* and 2 of 70

tests for *E. blennioides* ($P < 0.01$), indicating a lack of appreciable influence from null alleles, site-scale

Wahlund effects, or other violations of the Hardy-Weinberg model (Appendix H). Similarly, linkage

equilibrium was rejected in only 3 of 220 tests for *E. rufilineatum* and 3 of 210 tests for *E. blennioides*. I

therefore retained data from all sites and all loci for further analyses. The genetic diversity statistics H_E

and *A* were relatively high across sites for both species, though somewhat higher on average in *E. rufilineatum* than in *E. blennioides* (Table 4.1). In contrast, the population-stability index *M* generally was higher in *E. blennioides* than in *E. rufilineatum*. Several *E. rufilineatum* sites exhibited *M* values within the range (0.6-0.7) previously observed in populations known to have experienced recent bottlenecks (Garza and Williamson 2001). Between-site variability was similar between species for all three statistics (coefficient of variation range 0.03-0.10).

In several cases, low F_{ST} estimates between sites indicated that those sites belonged to the same population (Appendix J). For *E. rufilineatum*, three groups of sites exhibited no statistical evidence for departure from panmixia: (1) CLI1-CLI2, (2) LAUR-NFH1-NFH2, and (3) TEL1-TEL2. For *E. blennioides*, sites CLI1 and CLI2 were not differentiated. None of these groups transcended a major barrier to fish movement, but sites within groups were an average of 53 stream km apart (range 13-138 km). To avoid pseudoreplication, I retained for models of genetic diversity only the latter site in each group. Excluding within-population comparisons, pair-wise F_{ST} values were similar for the two species, averaging 0.06 (range 0.01-0.13) for *E. rufilineatum* and 0.04 (range -0.004 to 0.08) for *E. blennioides*. However, when these mean F_{ST} values were "standardized" by dividing by mean within-population homozygosity (Hedrick 2005), the overall degree of differentiation was substantially higher for *E. rufilineatum* than for *E. blennioides* (mean $G'_{ST} = 0.30$ and 0.16, respectively).

Random Forest models indicated the importance of stream size, site isolation, and urban land use for predicting genetic diversity of darter populations, but the strength and form of relationships varied among response variables and between species. Furthermore, partial dependence plots indicated that relationships often exhibited abrupt thresholds, with regressors influencing response variables only over a narrow range of values. Models for *E. rufilineatum* and *E. blennioides* explained 39% and 17%, respectively, of the variation in H_E among populations. For *E. rufilineatum*, H_E was most strongly related to site isolation; sites near a downstream reservoir exhibited lower diversity than sites far away from a reservoir (Figure 4.2). For both species, upstream watershed area, a measure of stream size, exhibited an apparent log-positive relationship with H_E (Figures 4.2 and 4.3).

Models for *E. rufilineatum* and *E. blennioides* explained 33% and 24%, respectively, of the variation in *A* among populations. For *E. rufilineatum*, *A* was positively related to watershed area, but negatively related to percentage of impervious surface present in HU12s (Figure 4.2). In the latter case, a sharp decrease in *A* occurred when imperviousness exceeded approximately 0.5%. For *E. blennioides*, *A* was negatively related to the percentage of developed land in HU12s and positively related to the distance to a downstream reservoir (Figure 4.3).

Models for *E. rufilineatum* and *E. blennioides* explained 47% and 12%, respectively, of the variation in *M* among populations. For *E. rufilineatum*, *M* was most strongly related to the density of roads and the percentage of impervious surfaces in HU12s (Figure 4.2). Values of *M* exhibited sharp declines when road density increased from 16 to 20 km ha^{-1} or impervious surfaces exceeded 1% of HU12 area. For *E. blennioides*, the only important predictor was percentage of developed land in HU12s, which exhibited a near-linear negative relationship with *M* (Figure 4.3).

Random Forest models indicated the importance of hydrologic alteration and, to a lesser extent stream size and spatial separation, in explaining patterns of genetic differentiation between sites. Models for *E. rufilineatum* and *E. blennioides* explained 26% and 11%, respectively, of the variation in F_{ST} among site-pairs. For *E. rufilineatum*, the best-supported index of hydrologic isolation discriminated among solely riverine connections, sites separated by a reservoir but not a dam, and sites separated by a dam. Increasing levels of hypothesized isolation within this ordinal scale corresponded with increasing mean values of F_{ST} (Figure 4.2). Other important variables for *E. rufilineatum* included the mean watershed area upstream of sites, which overall was negatively related to F_{ST}, and the fluvial distance separating sites, which overall was positively related to F_{ST}. However, these relationships were nuanced and exhibited thresholds. For example, the mean value of F_{ST} increased sharply with distance between 0 km and 100 km, but was slightly negatively related to distance over larger spatial extents. For *E. blennioides*, the only important predictor of F_{ST} was the presence or absence of an intervening dam, which resulted in a lower or higher mean value of F_{ST}, respectively (Figure 4.3).

DISCUSSION

Island biogeography of population genetic patterns in streams

Habitat patch size and connectivity influence population size, immigration rates, and extinction/colonization rates across a variety of ecosystem types (MacArthur and Wilson 1967), including temperate streams (reviewed in Roberts and Hitt 2010). These demographic phenomena can be measured by statistics calculated from variation at neutral genetic markers, often much more easily than by field-based methods (Lowe and Allendorf 2010; Luikart et al. 2010). Knowledge of relationships of patch size and connectivity to genetic diversity and differentiation is necessary in order to calibrate genetic statistics so that additive influences of land use or other anthropogenic impacts can be properly attributed. Moreover, such knowledge is relevant in its own right, by informing conservation biologists of: 1) the patch sizes necessary to maintain persistent populations, and 2) the magnitude of population isolation imposed by hypothesized dispersal barriers.

Not surprisingly, I found that major hydroelectric dams constituted substantial barriers to dispersal of both darter species, as indicated by high F_{ST} values between sites intervened by a dam. More surprising was my finding that pairs of *E. rufilineatum* sites located in adjacent tributaries of reservoirs exhibited F_{ST} values approaching those of site-pairs separated by dams. Most of these impoundments have been in place for approximately 10-25 darter generations, which evidently has been sufficient time for detectable genetic structure to develop. The influence of obvious movement barriers on population differentiation and diversity is well-studied for stream fishes (Hänfling and Weetman 2006; Neville et al. 2006; Pritchard et al. 2007; Bessert and Ortí 2008; Raeymaekers et al. 2008; Whiteley et al. 2010; see Chapter 1). In contrast, the influence of semi-permeable movement barriers, such low-quality habitat in lentic environments, is less studied. Skalski et al. (2008) found that gene flow of creek chub (*Semotilus atromaculatus*), a stream specialist, was significantly lower between tributaries of a reservoir than between tributaries of a free-flowing river. Reservoir-tributary populations also exhibited lower genetic diversity, potentially because these populations received less of a "rescue effect" from neighbor populations. Similarly, Mitchell et al. (2002) detected increased genetic differentiation among

126

populations of the lotic-specialized yellowcheek darter (*E. moorei*) located in adjacent tributaries of an impoundment. In contrast, Franssen (2012) found that the habitat generalist red shiner (*Cyprinella lutrensis*), which lives in both lentic and lotic habitats, exhibited high gene flow among sites regardless of whether they were separated by riverine or reservoir habitat.

Although *E. rufilineatum* occasionally has been observed in the littoral zone of reservoir coves in the UTRB (Etnier and Starnes 1993, page 459), the species is considered a lotic habitat specialist (Jenkins and Burkhead 1994). Accordingly, genetic differentiation data for this species indicated that reservoir habitat was more permeable than dams, but substantially less permeable than free-flowing habitat. An analogous test could not be conducted for *E. blennioides* because only one river-connected site pair was sampled; however, the F_{ST} estimated for this pair (0.004) was less than the average F_{ST} for sites separated by a reservoir (0.02) or dam (0.04) (Appendix J). Reservoir juxtaposition affected not only genetic differentiation, but diversity as well, given that sites closer to a downstream reservoir exhibited lower gene diversity (for *E. rufilineatum*) and allele richness (for *E. blennioides*) than sites farther from a reservoir. Apparently, proximity to a reservoir cuts a site off from downstream immigrants, which may be important for demographic supplementation and rescue (Gorman 1986; Winston et al. 2001).

Other correlates of population differentiation for *E. rufilineatum* included stream size and the distance between sites, although neither variable was as influential as hydrologic alteration. Presumably, larger streams housed larger darter populations, which correspondingly experienced slower drift, lower demographic stochasticity, and greater gene flow than populations occupying small tributaries (Koizumi et al. 2006; Raeymaekers et al. 2008). Interestingly, neither species exhibited a pronounced pattern of genetic isolation-by-distance (IBD) across the UTRB. *E. rufilineatum* exhibited IBD over only a narrow range of distances (0 to 100 km), while *E. blennioides* exhibited no evidence for IBD, once other variables were accounted for. Due to the correlation between distance and fragmentation (i.e., the farther away two sites are located, the more likely there is a dam between them), it is difficult to parse these two influences on differentiation. Likewise, it is difficult to draw inferences about patterns of pre-impoundment connectivity for these two species. However, it seems reasonable to assume that IBD was

the primary historical pattern of genetic variation among populations within the UTRB (e.g., Koizumi et al. 2006; Primmer et al. 2006; Whiteley et al. 2006). If so, this historical pattern has been overwhelmed by contemporary barriers to dispersal, producing a situation more closely resembling an isolation (Nei and Chakravarti 1977) or nonequilibrium metapopulation (Schlosser and Angermeier 1995) model of population structure (see also Chapter 1).

Another noteworthy finding was that, in the absence of a dam or reservoir barrier, neither species demonstrated substantial population structure. For example, pairs of *E. rufilineatum* sites located within the mainstem Clinch, North Fork Holston, and Tellico rivers exhibited F_{ST} values not different from zero, despite being separated by 26 to 138 km of stream. Maintenance of panmixia over such large spatial extents suggests that, in mainstem rivers, darters exhibit extensive migration, large effective population size, or both. In contrast, of the nine tributary-mainstem *E. rufilineatum* site-pairs not separated by a barrier, seven departed from panmixia (i.e., had F_{ST} values significantly greater than zero). This suggests that gene flow is higher within mainstems than between tributaries and mainstems, which is consistent with previously observed stream-size-related variation in population size, stability, and dispersal of fishes (reviewed in Roberts and Hitt 2010). The only *E. blennioides* comparison not separated by a barrier was between two mainstem Clinch River sites; as with *E. rufilineatum*, these sites exhibited panmixis over a large spatial extent.

The potentially large-grained population structure of these darter species contrasts with conventional wisdom that stream fish in general (Gerking 1953), and darters in particular (Schwalb et al. 2010), seldom move beyond stream-reach boundaries (see also Fausch et al. 2002 and Chapters 1 and 2). From an environmental assessment standpoint, this finding suggests that stream fishes integrate environmental conditions over large spatial scales, potentially across whole watersheds (see above). From a conservation standpoint, it suggests that monitoring and management activities need a regional focus on habitat connectivity and maintenance of ecosystem processes, rather than a narrow focus on local habitat protection and restoration (Schlosser and Angermeier 1995; Fullerton et al. 2010).

Influences of land use on population size and stability

Random Forest multiple regression models allowed me to test the influences of hypothesized land-use impacts on darter populations, after accounting for effects of stream size and fragmentation. These models implicated several land-use variables as being correlated with genetic indices of population size and stability, but the most consistently important predictors described variation among HU12 watersheds in the level of urbanization. Important urban land-use variables included % developed land by area, % impervious surfaces by area, and road density within HU12s. These variables were positively correlated with each other (Pearson's *r* ranged from 0.58 to 0.91), suggesting that they measured essentially the same phenomena. Watershed urbanization can increase fine-sediment deposition on the stream bottom, destabilize stream banks and channels, increase pollutant load, and increase the flashiness of streamflows (Allan 2004; Walsh et al. 2005), all of which could reduce habitat quality and quantity for benthic specialists like *E. rufilineatum* and *E. blennioides* (Peoples and Frimpong, in press). Although neither of these species is considered particularly sensitive to anthropogenic impacts, both showed reduced diversity and/or increased likelihood of past population instability as watershed urbanization increased. Such chronic influences may not be detectable from simple presence/absence or abundance measures of fish communities, which capture population status at only a snapshot in time. Furthermore, partial dependence plots showed that most of the decrease in genetic diversity occurred with very slight increases in watershed urbanization. These urbanization thresholds, often less than 1% of watershed area, are lower than previously published thresholds shown to predict extirpation of sensitive fish species (e.g., 2-12% imperviousness; Wenger et al. 2008) or shifts from intolerant to tolerant fish species assemblages (e.g., 8-12% imperviousness; Wang et al. 2001). This finding provides further evidence that genetic monitoring can provide an early warning of population decline, prior to a species' extirpation.

It is important to recognize that none of my regression models explained a large percentage of the variation in genetic statistics (range 12-47%), suggesting that other unmeasured factors contributed significantly to population fluctuations and/or variation in genetic estimates for non-demographic reasons (e.g., due to sampling or genotyping error). Furthermore, the small sample size of *E. blennioides* sites

renders models for that species tenuous and in need of further testing. However, overall findings were similar between the two species, in particular that stream size was positively related, and reservoir-related fragmentation and urban land use negatively related, to some measures of population size and stability.

It is interesting to note differences among the three genetic diversity statistics in the most important predictor variables. For both species, gene diversity (H_E) was related only to stream size and isolation, and the population stability index (M) was related only to land use, whereas allele richness (A) was related to a combination of these features. A is known to decrease faster and to reach a new equilibrium value sooner than H_E following a population size change, and thus can be viewed as the more contemporary indicator of population size (Schwartz et al. 2006). Thus, H_E in these darter populations may not yet have responded to recent land-use changes in the UTRB, but still primarily reflects the long-term signal of pre-impoundment variation in patch size. M, in contrast, is only weakly related to equilibrium effective population size, but is sensitive to *reductions* in population size (Garza and Williamson 2001). Therefore, M would not be expected to respond to island biogeographic factors such as stream size that deterministically drive long-term equilibrium population size, but would be expected to respond to contemporary phenomena such as anthropogenic impacts that cause stochastic reductions in population size. The combined use of these statistics provides a more complete picture of demographic history than would any statistic by itself, to an extent allowing inferences about the timing and duration of demographic impacts to populations.

Use of population-genetic markers to assess ecosystem condition

Although the use of genetic markers to infer biotic responses to ecosystem condition seems promising, several limitations bear further consideration. First, because fish populations may be geographically extensive (Albanese et al. 2004; Waits et al. 2008; see Chapter 1), it is difficult to draw direct conclusions about the influences of local environmental conditions on local population dynamics. This problem is not limited to population genetic approaches. Both individual- and community-based bioassessment techniques must account for the influences of regional conditions and immigrants on biotic metrics measured at local sites (e.g., Osborne and Wiley 1992; Larno et al. 2001; Hitt and Angermeier

2008). Genetic data allow dealing with this problem directly by delineating population structure and ensuring that environmental variables are measured over a spatial extent approximating that of population extents. In this study, I collected land-use data at the watershed (i.e., HU12) extent, which seemed to match darter population boundaries reasonably well: no HU12 contained multiple inferred populations, and few inferred populations transcended multiple HU12s. If one wished to draw inferences over narrower spatial extents, the best strategy would be to study species with more-limited dispersal capabilities and finer-grained population structure [e.g., mottled sculpin (*Cottus bairdi*); Lamphere and Blum 2012]. However, such a strategy might limit investigation to small streams, where sedentary species are most common (Woolnaugh et al. 2009).

Second, unlike toxicity assays performed on individuals or functional metrics calculated from communities, demographic estimates made from genetic data cannot be directly associated with particular stressors. Populations may decline or fluctuate due to a variety of individual or combined effects, such as floods, droughts, pollutant spills, predator or competitor introductions, chronic inputs of sediment or nutrients, or chance. Given that none of these impacts has a particular genetic "signature", population-genetic assessment metrics may characteristically have a lower signal:noise ratio than other types of metrics. This possibility certainly merits further study. However, demographic fluctuations are the ultimate determinants of species persistence, and the potential for their early detection (i.e., prior to extirpation) may outweigh the inability to pinpoint their cause.

Finally, population genetic data apply only to single species, and a focal species may not represent how other taxa respond to a given set of landscape conditions. Presumably, the use of multiple species (as in this study) and of species with contrasting ecological attributes strengthens inferences about biotic responses. Reliance on a particular species limits investigation to sites at which that species is present, in numbers great enough to obtain a reasonable sample-size of individuals. For this reason, common species are more likely than rare or endemic species to provide a wide range of land-use contrasts. Because certain common species are particularly valuable (e.g., for fishing), it may justify the cost of monitoring their genetic characteristics to ensure population persistence. However, common

131

species may be less sensitive than endemics to subtle land-use changes (Scott 2006), forcing a tradeoff between the extensiveness and sensitivity of a monitoring program. Future users of population genetic data to detect ecosystem condition should carefully select study species that occur across the region of interest, but are likely to exhibit demographic responses to the stressors of interest. Such selections could follow a rationale similar to that employed in community bioassessments methods, for example by monitoring genetic characteristics of suites of species that contrast in their habitat specialization, tolerance of stressors, and benthic versus pelagic habit.

ACKNOWLEDGMENTS

REFERENCES

Abell, R.A., and 10 other authors. 2000. Freshwater Ecoregions of North America: a Conservation Assessment. Island Press, Washington D.C.

Allan, J.D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology and Systematics 35:257-284.

Angermeier, P.L., K.L. Krueger, K.M. Mattson, and S.D. Klopfer. 2009. Development of an aquatic gap analysis for the upper Tennessee River basin. Project report to United States Geological Survey, Reston, VA.

Beneteau, C.L., N.E. Mandrak, and D.D. Heath. 2009. The effects of river barriers and range expansion on the population genetic structure and stability in greenside darter (*Etheostoma blennioides*) populations. Conservation Genetics 10:477-487.

Berkman, H.E., and C.F. Rabeni. 1987. Effect of siltation on stream fish communities. Environmental Biology of Fishes 18:285-294.

Bessert, M.L., and G. Ortí. 2008. Genetic effects of habitat fragmentation on blue sucker populations in the upper Missouri River (*Cycleptus elongatus* Lesueur, 1918). Conservation Genetics 9:821-832.

Breiman, L. 2001. Random forests. Machine Learning 45:5-32.

Bunn, S.E., and J.M. Hughes. 1997. Dispersal and recruitment in streams: evidence from genetic studies. Journal of the North American Benthological Society 16:338-346.

Cutler, D.R., T.C. Edwards, Jr., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler. 2007. Random forests for classification in ecology. Ecology 88:2783-2792.

De'ath, G., and K.E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

DeWoody, J.A., D.E. Fletcher, S.D. Wilkins, and J.C. Avise. 2000. Parentage and nest guarding in the tessellated darter (*Etheostoma olmstedi*) assayed by microsatellite markers (Perciformes: Percidae). Copeia 2000:740-747.

Diamond, J.M., D.W. Bressler, and V.B. Serveiss. 2002. Assessing relationships between human land uses and the decline of native mussels, fish and macroinvertebrates in the Clinch and Powell river watershed, USA. Environmental Toxicology and Chemistry 21:1147-1155.

Dunn, T., and L.B. Leopold. 1978. Water in Environmental Planning. W.H. Freeman, San Francisco, CA.

Etnier, D.A. 1997. Jeopardized southeastern freshwater fishes: a search for causes. Pages 87-92 in G.E. Benz and D.E. Collins (editors), Aquatic Fauna in Peril: The Southeastern Perspective. Lenz Design and Communications, Decatur, GA.

Etnier, D.A., and W.C. Starnes. 1993. The Fishes of Tennessee. University of Tennessee Press, Knoxville, Tennessee.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin version 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Fausch K.D., C.E. Torgersen, C.V. Baxter, and H.W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52:483-498.

Franssen, N.R. 2012. Genetic structure of a native cyprinid in a reservoir-altered stream network. Freshwater Biology 57:155-165.

Friedman, J.H. 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics 29:1189-1232.

Fullerton, A.H., and eight other authors. 2010. Hydrological connectivity for riverine fish: measurement challenges and research opportunities. Freshwater Biology 55:2215-2237.

Gabel, J.M., E.E. Dakin, B.J. Freeman, and B.A. Porter. 2008. Isolation and identification of eight microsatellite loci in the Cherokee darter (*Etheostoma scotti*) and their variability in other members of the genera *Etheostoma*, *Ammocrypta*, and *Percina*. Molecular Ecology Resources 8:149-151.

Garza, J.C., and E.G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. Molecular Ecology 10:305-318.

Gerking, S.D. 1953. Evidence for the concepts of home range and territory in stream fishes. Ecology 34:347-365.

Gorman, O.T. 1986. Assemblage organization of stream fishes: the effect of rivers on adventitious streams. The American Naturalist 128:611-616.

Gotelli, N.J., and C.M. Taylor. 1999. Testing metapopulation models with stream-fish assemblages. Evolutionary Ecology Research 1:835-845.

Goudet, J. 2002. FSTAT: a program to estimate and test gene diversity and fixation indices. Available online at http://www.unil.ch/izea/softwares/fstat.html.

Hampson, P.S., M.W. Treece, G.C. Johnson, S.A. Ahlstedt, and J.F. Connell. 2000. Water quality in the upper Tennessee River basin, Tennessee, North Carolina, Virginia, and Georgia 1994-98. U.S. Geological Survey Circular 1205.

Hänfling, B., and D. Weetman. 2006. Concordant genetic estimators of migration reveal

    anthropogenically enhanced source-sink population structure in the river sculpin, *Cottus gobio*.

    Genetics 173:1487-1501.

Harding, J.S., E.F. Benfield, P.V. Bolstad, G.S. Helfman, and E.B.D. Jones III. 1998. Stream

    biodiversity: the ghost of land use past. Proceedings of the National Academy of Science USA

    95:14843-14847.

Heithaus, M.R., and R.H. Laushman. 1997. Genetic variation and conservation of stream fishes:

    influence of ecology, life history, and water quality. Canadian Journal of Fisheries and Aquatic

    Sciences 54:1822-1836.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.

Hilderbrand, R.H., and J.L. Kershner. 2000. Conserving inland cutthroat trout in small streams: how

    much stream is enough? North American Journal of Fisheries Management 20:513-520.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family-level biotic index.

    Journal of the North American Benthological Society 7:65-68.

Hitt, N.P., and P.L. Angermeier. 2008. Evidence for fish dispersal from spatial analysis of stream

    network topology. Journal of the North American Benthological Society 27:304-320.

Horwitz, R.J. 1978. Temporal variability patterns and the distributional patterns of stream fishes.

    Ecological Monographs 48:307-321.

Jelks, H.L., and 15 co-authors. 2008. Conservation status of imperiled North American freshwater and

    diadromous fishes. Fisheries 33:372-407.

Jenkins, R.E., and N.M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society,

    Bethesda, Maryland.

Jones, J.A., F.J. Swanson, B.C. Wemple, and K.U. Snyder. 2000. Effects of roads on hydrology,

    geomorphology, and disturbance patches in stream networks. Conservation Biology 14:76-85.

Junk, W.J., P.B. Bayley, and R.E. Sparks. 1989. The flood pulse concept in river-floodplain systems.

    Canadian Special Publications in Fisheries and Aquatic Sciences 106:110-127.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey Special Publication 5.

Koizumi, I., S. Yamamoto, and K. Maekawa. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden char. Molecular Ecology 15:3175-3189.

Lamphere, B.A., and M.J. Blum. 2012. Genetic estimates of population structure and dispersal in a benthic stream fish. Ecology of Freshwater Fish 21:75-86.

Larno, V., J. Laroche, S. Launey, P. Flammarion, and A. Devaux. 2001. Responses of chub (*Leuciscus cephalus*) populations to chemical stress, assessed by genetic markers, DNA damage, and cytochrome P4501A induction. Ecotoxicology 10:145-158.

Lenat, D.R. 1984. Agriculture and stream water quality – a biological evaluation of erosion control practices. Environmental Management 23:257-270.

Lowe, W.H., and F.W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051.

Luikart, G., N. Ryman, D.A. Tallmon, M.K. Schwartz, and F.W. Allendorf. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conservation Genetics 11:355-373.

MacArthur, R.H., and E.O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, NJ.

Mattson, K.M., and P.L. Angermeier. 2007. Integrating human impacts and ecological integrity into a risk-based protocol for conservation planning. Environmental Management 39:125-138.

Mitchell, R.M., R.L. Johnson, and G.L. Harp. 2002. Population structure of an endemic species of yellowcheek darter, *Etheostoma moorei* (Raney and Suttkus), of the upper Little Red River, Arkansas. American Midland Naturalist 148:129-137.

Morita, K., and S. Yamamoto. 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. Conservation Biology 16:1318-1323.

Nei, M., and A. Chakravarti. 1977. Drift variances of F_{ST} and G_{ST} statistics obtained from a finite number of isolated populations. Theoretical Population Biology 11:307-325.

Neves, R.J., and P.L. Angermeier. 1990. Habitat alteration and its effects on native fishes in the upper Tennessee River system, east-central USA. Journal of Fish Biology 37:45-52.

Neville, H.M., J.B. Dunham, and M.M. Peacock. 2006. Landscape attributes and life history variability shape genetic structure of trout populations in a stream network. Landscape Ecology 21:901-916.

Osborne, L.L., and M.J. Wiley. 1988. Empirical relationships between land use/cover and stream water quality in an agricultural watershed. Journal of Environmental Management 26:9-27.

Osborne, L.L., and M.J. Wiley. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. Canadian Journal of Fisheries and Aquatic Sciences 49:671-681.

Peoples, B.K., and E.A. Frimpong. In press. A snapshot analysis of age distributions of fishes in urban and forested Virginia streams. Urban Ecosystems.

Poff, N.L., and J.V. Ward. 1989. Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow patterns. Canadian Journal of Fisheries and Aquatic Sciences 46:1805-1818.

Poff, N.L., and seven other authors. 1997. The natural flow regime: a paradigm for river conservation and restoration. BioScience 47:769-784.

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

Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9:181-199.

Primmer, C.R., A.J. Veselov, A. Zubchenko, A. Poututkin, I. Bakhmet, and M.T. Koskinen. 2006. Isolation by distance within a river system: genetic population structuring of Atlantic salmon,

Salmo salar, in tributaries of the Varzuga River in northwest Russia. Molecular Ecology 15:653-666.

Pritchard, V.L., K. Jones, and D.E. Cowley. 2007. Genetic diversity within fragmented cutthroat trout populations. Transactions of the American Fisheries Society 136:606-623.

R Development Core Team. 2004. *R*: a language and environment for statistical computing. Available online at http://www.R-project.org.

Raeymaekers, J.A.M., G.E. Maes, S. Geldof, I. Hontis, K. Nackaerts, and F.A.M. Volckaert. 2008. Modeling genetic connectivity in sticklebacks as a guideline for river restoration. Evolutionary Applications 1:475-488.

Roberts, J.H., and P.L. Angermeier. 2007. Spatiotemporal variability of stream habitat and movement of three species of fish. Oecologia 151:417-430.

Roberts, J.H., and N.P. Hitt. 2010. Longitudinal structure in temperate stream fish communities: evaluating conceptual models with temporal data. American Fisheries Society Symposium 73:281-299.

Schlosser, I.J. 1987. A conceptual framework for fish communities in small warmwater streams. Pages 17-24 in W. J. Matthews and D. C. Heins (editors). Community and Evolutionary Ecology of North American Stream Fishes. Oklahoma University Press, Norman, OK.

Schlosser, I.J., and P.L. Angermeier. 1995. Spatial variation in demographic processes of lotic fishes: conceptual models, empirical evidence, and implications for conservation. American Fisheries Society Symposium 17:392-401.

Schwalb, A.N., M.S. Poos, and J.D. Ackerman. 2010. Movement of logperch – the obligate host fish for endangered snuffbox mussels: implications to mussel dispersal. Aquatic Sciences 73:223-231.

Schwartz, M.K., G. Luikart, and R.S. Waples. 2006. Genetic monitoring as a promising tool for conservation and management. Trends in Ecology and Evolution 22:25-33.

Scott, M.C. 2006. Winners and losers among stream fishes in relation to land use legacies and urban development in the southeastern US. Biological Conservation 127:301-309.

Sheldon, A.L. 1988. Conservation of stream fishes: patterns of diversity, rarity, and risk. Conservation Biology 2:149-156.

Skalski, G.T., J.B. Landis, M.J. Grose, and S.P. Hudman. 2008. Genetic structure of creek chub, a headwater minnow, in an impounded river system. Transactions of the American Fisheries Society 137:962-975.

Sugg, D.W., R.K. Chesser, J.A. Brooks, and B.T. Grasman. 1995. The association of DNA damage to concentrations of mercury and radiocesium in largemouth bass. Environmental Toxicology and Chemistry 14:661-668.

Sullivan, K.B., and M.J. Lydy. 1999. Differences in survival functions of mosquitofish (*Gambusia affinis*) and sand shiner (*Notropis ludibundus*) genotypes exposed to pesticides. Environmental Toxicology and Chemistry 18:906-911.

Taylor, C.M., and M.L. Warren. 2001. Dynamics in species composition of stream fish assemblages: environmental variability and nested subsets. Ecology 82:2320-2330.

Thompson, W.L., G.C. White, and C. Gowan. 1998. Monitoring Vertebrate Populations. Academic Press, San Diego, CA. 365 pp.

Townsend, C.R., and A.G. Hildrew. 1994. Species traits in relation to a habitat templet for river systems. Freshwater Biology 31:265-275.

Van Straalen, N.M., and M.J.T.N. Timmermans. 2002. Genetic variation in toxicant-stressed populations: an evaluation of the "genetic erosion" hypothesis. Human and Ecological Risk Assessment 8:983-1002.

Vannote, R.L., W.G. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130-137.

Vrijenhoek, R.C., E. Pfeiler, and J.D. Wetherington. 1992. Balancing selection in a desert stream-dwelling fish, *Poeciliopsis monacha*. Evolution 46:1642-1657.

Waits, E.R., M.J. Bagley, M.J. Blum, F.H. McCormick, and J.M. Lazorchak. 2008. Source-sink dynamics sustain central stonerollers (*Campostoma anomalum*) in a heavily urbanized catchment. Freshwater Biology 53:2061-2075.

Walsh, C.J., A.H. Roy, J.W. Feminella, P.D. Cottingham, P.M. Groffman, and R.P. Morgan II. 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society. 24:706-723.

Wang, L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28:255-266.

Wear, D.N., and P. Bolstad. 1998. Land-use changes in southern Appalachian landscapes: spatial analysis and forecast evaluation. Ecosystems 1:575-594.

Weir, B.S., and C.C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358-1370.

Wenger, S.J., J.T. Peterson, M.C. Freeman, B.J. Freeman, and D.D. Homans. 2008. Stream fish occurrence in response to impervious cover, historic land use, and hydrogeomorphic factors. Canadian Journal of Fisheries and Aquatic Sciences 65:1250-1264.

Whiteley, A.R., P. Spruell, and F.W. Allendorf. 2006. Can common species provide valuable information for conservation? Molecular Ecology 15:2767-2786.

Whiteley, A.R., K. Hastings, J.K. Wenburg, C.A. Frissell, J.C. Martin, and F.W. Allendorf. 2010. Genetic variation and effective population size in isolated populations of coastal cutthroat trout. Conservation Genetics 11:1929-1943.

Winn, H.E. 1958. Comparative reproductive behavior and ecology of fourteen species of darters (Pisces-Percidae). Ecological Monographs 28: 155-191.

Winston, M.R., C.M. Taylor, and J. Pigg. 1991. Upstream extirpation of four minnow species due to damming of a prairie stream. Transactions of the American Fisheries Society 120:98-105.

Woolnaugh, D.A., J.A. Downing, and T.J. Newton. 2009. Fish movement and habitat use depends on water body size and shape. Ecology of Freshwater Fish 18:83-91.

Table 4.1. Characteristics of 23 sites sampled for *Etheostoma rufilineatum* and *E. blennioides*. Site locations are depicted in Figure 4.1.

Site code	Site waterbody	Receiving waterbody	Latitude	Longitude	Upstream watershed area (km^2)	Distance to downstream reservoir (km)
BIGC	Big Creek	John Sevier Detention Reservoir	36.418	-82.952	134.8	1.1
BMOC	Big Moccasin Creek	North Fork Holston River	36.676	-82.529	190.3	76.0
BSYC	Big Sycamore Creek	Norris Lake	36.450	-83.440	56.1	7.7
BULL	Bullrun Creek	Melton Hill Lake	36.163	-83.946	148.3	24.4
CLI1	Clinch River downstream	Norris Lake	36.580	-83.005	3080.0	64.1
CLI2	Clinch River upstream	Norris Lake	36.964	-82.076	1284.0	202.5
COPP	Copper Creek	Clinch River	36.735	-82.443	161.2	148.9
EMOR	Emory River	Watts Bar Lake	36.027	-84.579	1822.7	11.4
FLAT	Flat Creek	Holston River	36.078	-83.744	173.4	3.5
LAUR	Laurel Creek	North Fork Holston River	36.923	-81.673	177.0	198.9
LICK	Lick Creek	Nolichucky River	36.152	-83.136	678.4	25.4
LITT	Little River	Fort Loudon Lake	35.765	-83.855	498.5	21.4
LPIG	Little Pigeon River	French Broad River	35.816	-83.436	211.1	24.1
MFHR	Middle Fork Holston River	South Holston Lake	36.784	-81.697	458.6	41.9
NFH1	North Fork Holston River downstream	Holston River	36.790	-82.027	1023.9	148.7
NFH2	North Fork Holston River upstream	Holston River	36.898	-81.746	598.1	185.7
NOLI	Nolichucky River	Douglas Lake	36.099	-83.053	3341.9	38.6
POWE	Powell River	Norris Lake	36.621	-83.285	1185.2	110.3
SFHR	South Fork Holston River	South Holston Lake	36.654	-81.888	884.8	0.6
TEL1	Tellico River downstream	Tellico Lake	35.417	-84.259	358.4	8.4
TEL2	Tellico River upstream	Tellico Lake	35.325	-84.178	189.6	34.7
WACK	Wallen Creek	Powell River	36.644	-83.076	84.1	150.8
WHIT	Whites Creek	Watts Bar Lake	35.805	-84.769	317.4	3.8

Table 4.2. Genetic statistics for *Etheostoma rufilineatum* and *E. blennioides* sampled at 23 sites in the upper Tennessee River basin. Site codes correspond to Table 4.1 and Figure 4.1. Entries indicate the sample size of individuals (n), unbiased gene diversity (H_E), allele richness (A), and bottleneck index (M) at each site. The mean and coefficient of variation (CV) of each statistic across sites are shown. Dashes indicate sites not sampled for a given species.

| Site code | *E. rufilineatum* | | | | *E. blennioides* | | | |
	n	H_E	A	M	n	H_E	A	M
BIGC	24	0.762	8.4	0.768	-	-	-	-
BMOC	23	0.731	7.7	0.601	-	-	-	-
BSYC	23	0.761	7.9	0.731	-	-	-	-
BULL	24	0.834	8.8	0.601	-	-	-	-
CLI1	24	0.843	9.8	0.757	25	0.746	8.0	0.836
CLI2	24	0.845	9.2	0.720	13	0.746	9.0	0.760
COPP	23	0.811	9.2	0.714	-	-	-	-
EMOR	24	0.823	9.8	0.640	24	0.730	7.6	0.832
FLAT	23	0.739	6.7	0.707	-	-	-	-
LAUR	24	0.799	7.4	0.797	-	-	-	-
LICK	24	0.818	8.5	0.604	-	-	-	-
LITT	24	0.796	9.0	0.708	25	0.720	7.7	0.727
LPIG	24	0.837	9.9	0.667	24	0.681	6.9	0.872
MFHR	23	0.842	8.2	0.674	-	-	-	-
NFH1	24	0.806	8.8	0.712	-	-	-	-
NFH2	24	0.797	7.5	0.726	20	0.715	7.6	0.809
NOLI	-	-	-	-	25	0.720	7.3	0.872
POWE	24	0.849	9.2	0.806	12	0.763	8.9	0.786
SFHR	24	0.844	9.9	0.728	-	-	-	-
TEL1	24	0.735	8.5	0.764	-	-	-	-
TEL2	24	0.719	7.8	0.777	13	0.710	7.2	0.831
WACK	24	0.814	8.6	0.792	-	-	-	-
WHIT	24	0.826	8.4	0.658	23	0.700	7.0	0.702
Mean		0.801	8.6	0.711		0.723	7.7	0.803
CV		0.05	0.10	0.09		0.03	0.09	0.07
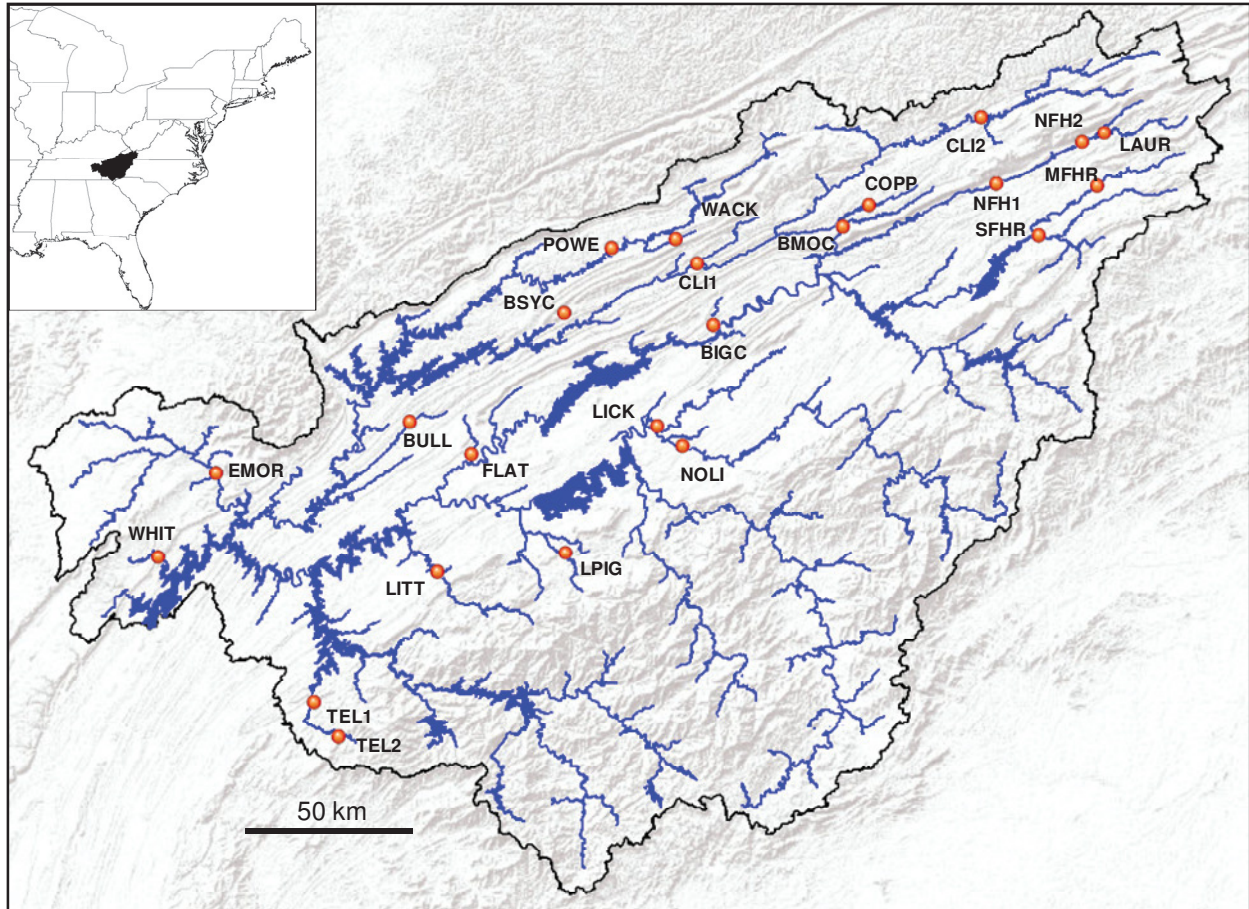
Figure 4.1. Map of the upper Tennessee River basin (UTRB), showing locations of sites (red circles) sampled for *Etheostoma rufilineatum* and *E. blennioides*. Site codes correspond to those presented in Table 4.1. Inset shows location of the UTRB (shaded area) within the eastern United States.
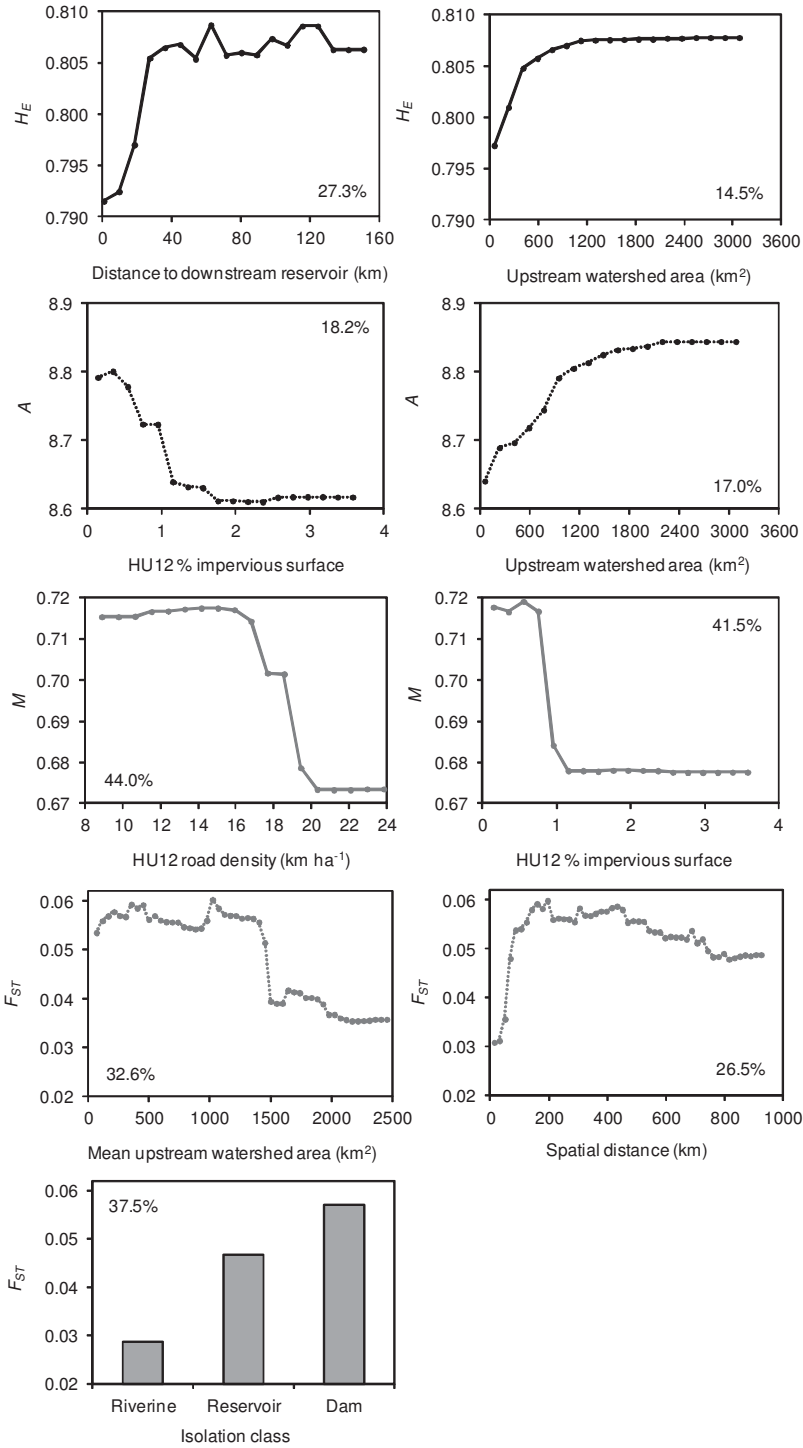
Figure 4.2. Partial dependence plots for regressors with importance scores >10% in Random Forest models for *Etheostoma rufilineatum*. Plots illustrate the effect of varying levels of a regressor on the mean of a response variable when all other regressors in the model are averaged out. Each plot shows the importance score of a regressor, the percentage increase in model error when that regressor is randomized among observations. Variables are described in greater detail in the text.

144
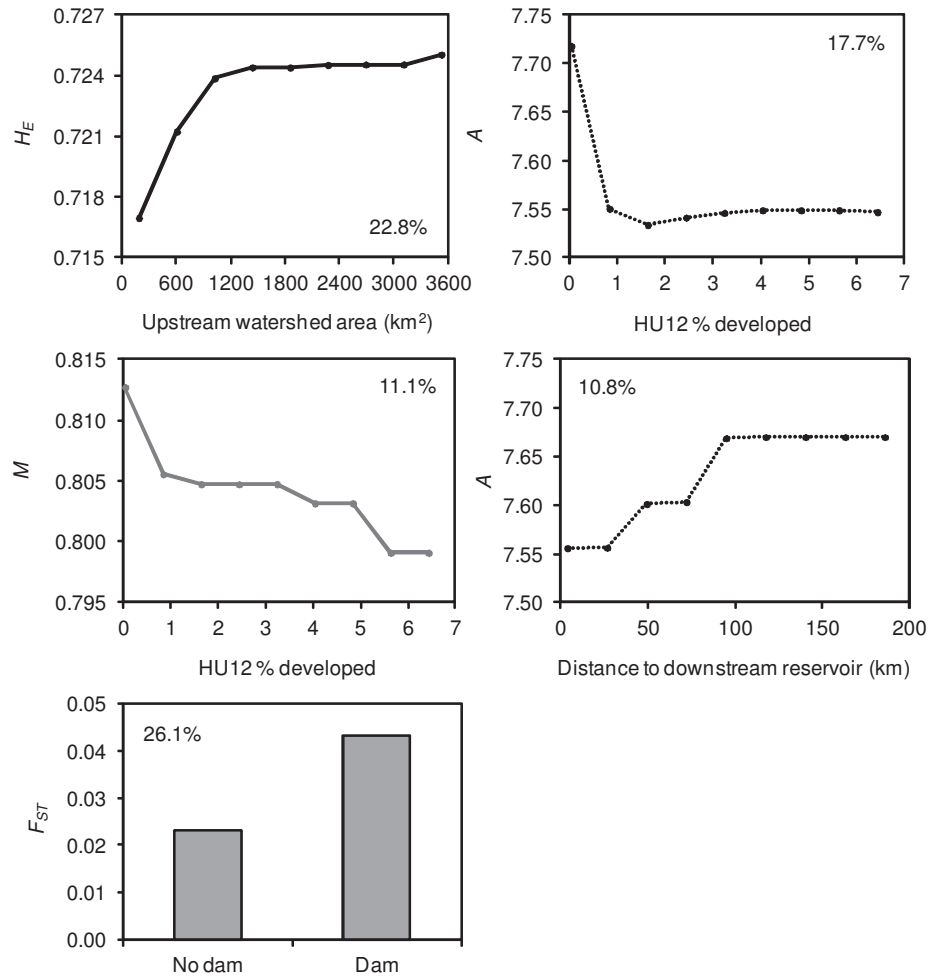
Figure 4.3. Partial dependence plots for regressors with importance scores >10% in Random Forest models for *Etheostoma blennioides*. Plots illustrate the effect of varying levels of a regressor on the mean of a response variable when all other regressors in the model are averaged out. Each plot shows the importance score of a regressor, the percentage increase in model error when that regressor is randomized among observations. Variables are described in greater detail in the text.

GENERAL CONCLUSIONS

I undertook this dissertation in order to better understand the ecology and evolution of three species of darters, and, where appropriate, generalize these findings to other stream fishes. The overall goals of my dissertation were to improve scientific understanding of 1) the spatial grains and extents over which darter populations and metapopulations are organized, 2) how darter populations respond to anthropogenic modifications of the landscape, 3) how genetic and ecological data can be used to assess the history of and predict future risks to darter populations, and 4) how management guidelines can be prescribed based on this knowledge. In the following sections, I summarize what I have learned along each of these lines of investigation.

Spatial scaling of population and metapopulation processes

Stream fishes have been the focus of a long-standing controversy in the ecological literature. Early mark-recapture studies suggested that many fish species carry out their lifecycles within a single stream-reach (Gerking 1953; Hill and Grossman 1987), a finding that ultimately codified the so-called "restricted movement paradigm" of stream-fish movement (Gowan et al. 1994). The results of such studies have been questioned on the basis of methodological flaws that downwardly bias estimates of movement (Albanese et al. 2003). However, I propose that the paradigm persists, because most monitoring and restoration initiatives for stream fishes still are undertaken over small spatial extents (i.e., tens to hundreds of meters) (e.g., Meador et al. 1993; Bernhardt et al. 2005), presumably because of the implicit assumption that these small extents capture key population processes. Yet if population or metapopulation dynamics play out over greater spatial extents, influences of local habitat conditions on fish abundance may be overwhelmed by regional immigration-emigration dynamics and measures of local abundance may have little utility for assessing population status (Gowan and Fausch 1996; Fausch et al. 2002; Hitt and Angermeier 2008).

My population genetic studies of *Percina rex*, *Etheostoma rufilineatum*, and *E. blennioides* indicate that the population dynamics of these species play out over large spatial extents, including entire streams and watersheds. Plots of genetic isolation-by-distance generally did not show an increase in slope

until populations were separated by more than 80-100 km, providing indirect evidence that darters underwent high gene flow at such spatial extents. Furthermore, based on direct methods, I estimated that juvenile *P. rex* commonly dispersed >14 km and occasionally dispersed up to 57 km within the upper Roanoke River watershed. Genetic panmixia over these spatial extents suggests that watersheds should be considered "patchy-populations" that exhibit frequent between-patch dispersal (Schlosser and Angermeier 1995; Falke and Fausch 2010). Management for these and potentially other darter species should be targeted at entire watersheds, embracing a spatial focus much more extensive than that adopted in many previous cases (Fausch et al. 2002).

Dispersal over extents greater than 100 km generally was prevented not by distance *per se*, but by impassible barriers such as dams and habitat conditions made unsuitable by anthropogenic activities. For example, *P. rex* showed no evidence for dispersal among two populations that were separated by unstable tailwater conditions, whereas *E. rufilineatum* populations were fragmented by lentic reservoir conditions. Thus, dispersal of these species was "all-or-nothing": high within populations but low to absent between populations. The resulting contemporary population structures of these species are best described by a nonequilibrium metapopulation demographic model (Schlosser and Angermeier 1995) or an isolation (Nei and Chakravarti 1977) evolutionary model. The demographic and evolutionary consequences of this new, fragmented condition merit additional study. Ultimately, conservation of stream-fish species may require the reunification of metapopulations, either through barrier removal or through intentional translocation of individuals among populations.

Responses of darter populations to anthropogenic modifications of the landscape

As described above, populations of all three species that I studied were highly fragmented by dams and associated reservoirs and tailwaters. Because these species are benthic specialists on clean substrate, their avoidance of crossing silted lacustrine habitats came as no surprise. Less investigated, and perhaps more interesting, was the degree to which reservoir-isolated populations suffered reduced genetic diversity and lowered effective population size, relative to populations connected to more riverine habitats. For *P. rex*, three populations presently are restricted by such unsuitable habitats to small

147

geographic ranges that span only one or two stream channels. These populations exhibit small census and effective population sizes and evidence for past population bottlenecks, and accordingly are at elevated risk of extinction. Similarly, *E. rufilineatum* populations occupying small streams that feed directly into a reservoir exhibited lower heterozygosity than those occupying larger streams that feed into other streams. This shows that even a small-bodied, locally abundant species like *E. rufilineatum* may require large spatial extents to complete its life cycle and/or may rely on immigrants to maintain long-term local persistence.

Regression models indicated that watershed urbanization reduced the size and stability of *E. rufilineatum* and *E. blennioides* populations in the upper Tennessee River basin. Urban land-use variables such as percentage developed land, percentage impervious surfaces, and road density in the watershed surrounding sampling sites were negatively correlated with genetic diversity and population stability index, two surrogates for demographic history. Urbanization increases sediment and pollutant loading and makes stream flows more erratic, all factors that could decrease habitat suitability and stability for stream fishes (Allan 2004). Interestingly, the urbanization thresholds (e.g., 1% imperviousness) at which I detected genetic responses were considerably lower than thresholds at which other studies have first detected responses in other biological variables (e.g., loss of sensitive species, changes in community composition; Wang et al. 2001; Wenger et al. 2008). This suggests that genetic monitoring of population status could provide an early warning of ecological impacts and a useful complement to existing biomonitoring protocols.

Assessing the history of and predicting future risks to populations

One of the main allures of genetic analyses is the ability to infer historical demographic and evolutionary events that cannot be understood using contemporary field studies (Schwartz et al. 1998). I had mixed success inferring these histories for darters. One of the main limitations is that historical patterns of population size and gene flow are easily masked by contemporary bottlenecks and fragmentation events that rapidly inflate drift. For example, I could not confidently describe evolutionary relationships between *P. rex* populations, because most populations have lost a substantial component of

148

their genetic diversity and are now fixed for a small, often non-overlapping set of alleles. I presume that historical migration followed an isolation-by-distance pattern that has been masked by contemporary fragmentation and non-spatial variation in drift. However, there is no definitive test of this hypothesis. Similarly, although approximate Bayesian computation models for *P. rex* allowed me to estimate with reasonable precision the contemporary effective population size and the number of generations back at which fragmentation and bottlenecks occurred, I could not "see past the event" and draw precise inferences about pre-fragmentation population sizes or migration rates.

As with any methodology that seeks to infer process from pattern, historical genetic inferences can break down when a given set of data are consistent with a wide range of demographic scenarios. This limitation ultimately may be remedied by likelihood-based modeling frameworks that extract additional information from genetic data (e.g., Beerli and Felsenstein 2001), but presently such frameworks are limited to simplified demographies that poorly match the presumed histories of my study species. Given the potential for drift to inflate the divergence of *P. rex* populations, the best I could do was estimate the *maximum* genetic divergence and time-since-isolation of these populations. Across various forms of evidence and in support of previous claims (Jenkins and Burkhead 1994), this divergence was relatively shallow and indicative of high historical connectivity among populations, perhaps as recently as the early 20th century.

Although genetic data had limited power to infer the history of darter populations, they proved useful for assessing future risks. I developed genetic-based (as well as demographic- and geographic-based) metrics of risk to populations incurred by adopting either of two management strategies: 1) small-population risks incurred by maintaining isolation, and 2) outbreeding-depression risks incurred by translocating fish among divergent populations. I then developed a set of risk criteria for *P. rex*, applied the criteria to calculated values of metrics, and developed aggregate measures of risk under each management option. Based on the results, several populations of *P. rex* could potentially benefit from genetic rescue to prevent inbreeding depression and lower the risk of extinction, whereas at least one of the other populations exhibits enough divergence that its continued isolation may be prudent. This

149

general framework for risk-assessment could be refined through better understanding of quantitative relationships between risk metrics and persistence. However, even in its present form, the framework could be a useful tool for evaluating and communicating risk and prioritizing management actions for fragmented populations of fish and other organisms.

REFERENCES

Albanese, B.W., P.L. Angermeier, and C. Gowan. 2003. Designing mark-recapture studies to reduce effects of distance weighting on movement distance distributions of stream fishes. Transactions of the American Fisheries Society 132:925-939.

Allan, J.D. 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annual Review of Ecology and Systematics 35:257-284.

Beerli, P., and J. Felsenstein. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations using a coalescent approach. Proceedings of the National Academy of Sciences USA 98:4563-4568.

Bernhardt, E.S., and 24 other authors. 2005. Synthesizing U.S. river restoration efforts. Science 308:636-637.

Caughley, G. 1994. Directions in conservation biology. Journal of Animal Ecology 63:215-244.

Fagan, W.F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. Ecology 83:3243-3249.

Falke, J.A., and K.D. Fausch. 2010. From metapopulations to metacommunities: linking theory with empirical observations of the spatial population dynamics of stream fishes. American Fisheries Society Symposium 73:207-233.

Fausch, K.D., C.E. Torgersen, C.V. Baxter, and H.W. Li. 2002. Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. BioScience 52:483-498.

Frankham, R., J.D. Ballou, and D.A. Briscoe. 2010. Introduction to Conservation Genetics, second edition. Cambridge University Press, Cambridge, U.K.

Frankham, R., and six other authors. 2011. Predicting the probability of outbreeding depression. Conservation Biology 25:465-475.

Gerking, S.D. 1953. Evidence for the concepts of home range and territory in stream fishes. Ecology 34:347-365.

Gowan, C., M.K. Young, K.D. Fausch, and S.C. Riley. 1994. Restricted movement in resident stream salmonids: a paradigm lost? Canadian Journal of Fisheries and Aquatic Sciences 51:2626-2637.

Gowan, C., and K.D. Fausch. 1996. Long-term demographic responses of trout populations to habitat manipulation in six Colorado streams. Ecological Applications 6:931-946.

Hill, J., and G.D. Grossman. 1987. Home range estimates for three North American stream fishes. Copeia 1987:376-380.

Hitt, N.P. and P.L. Angermeier. 2008. Evidence for fish dispersal from spatial analysis of stream network topology. Journal of the North American Benthological Society 27:304-320.

Jenkins, R.E. and N.M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Koizumi, I., S. Yamamoto, and K. Maekawa. 2006. Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden char. Molecular Ecology 15:3175-3189.

Lowe, W.H., and F.W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19:3038-3051.

McElhany, P., M.H. Ruckelshaus, M.J. Ford, T.C. Wainwright, and E.P. Bjorkstedt. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. United States Department of Commerce, NOAA Technical Memorandum NMFS-NWFSC-42.

Meador, M.R., T.F. Cuffney, and M.E. Gurtz. 1993. Methods for sampling fish communities as part of the National Water-Quality Assessment Program. United States Geological Survey, Open-File Report 93-104. Raleigh, North Carolina.

Morita, K. and S. Yamamoto. 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. Conservation Biology 16:1318-1323.

Nei, M. and A. Chakravarti. 1977. Drift variances of F_{ST} and G_{ST} statistics obtained from a finite number of isolated populations. Theoretical Population Biology 11:307-325.

Quinn, T.P., M.T. Kinnison, and M.J. Unwin. 2001. Evolution of Chinook salmon (*Oncorhynchus tshawytscha*) populations in New Zealand: pattern, rate, and process. Genetica 112:493-513.

Reiman, B.E., and J.B. Dunham. 2000. Metapopulations and salmonids: a synthesis of life history patterns and empirical observations. Ecology of Freshwater Fish 9:51-64.

Ricciardi, A. and J.B. Rasmussen. 1999. Extinction rates of North American freshwater fauna. Conservation Biology 13:1220-1222.

Roberts, J.H. and P.L. Angermeier. 2007. Spatiotemporal variability of stream habitat and movement of three species of fish. Oecologia 151:417-430.

Roberts, J.H. and A.E. Rosenberger. 2008. Threatened fishes of the world: *Percina rex* (Jordan and Evermann 1889) (Percidae). Environmental Biology of Fishes 83:439-440.

Schlosser, I.J. and P.L. Angermeier. 1995. Spatial variation in demographic processes of lotic fishes: conceptual models, empirical evidence, and implications for conservation. American Fisheries Society Symposium 17:392-401.

Schwartz, M.K., D.A. Tallmon, and G. Luikart. 1998. Review of DNA-based census and effective population size estimators. Animal Conservation 1:293-299.

Wang, L., J. Lyons, P. Kanehl, and R. Bannerman. 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28:255-266.

Waples, R.S. 1995. Evolutionarily significant units and the conservation of biological diversity under the Endangered Species Act. American Fisheries Society Symposium 17:8-27.

Wenger, S.J., J.T. Peterson, M.C. Freeman, B.J. Freeman, and D.D. Homans. 2008. Stream fish occurrence in response to impervious cover, historic land use, and hydrogeomorphic factors. Canadian Journal of Fisheries and Aquatic Sciences 65:1250-1264.

APPENDIX A: Results of Bayesian clustering models for *Percina rex*

Table A1. Comparison of STRUCTURE models with varying hypothesized numbers of genetic clusters (*K*), for all 578 *Percina rex* individuals combined and for each population. Entries indicate a model's loss of information (ΔAIC_c) relative to the best model (in bold). Population codes follow those of Table 1.1.

K	All individuals ΔAIC_c	UROAN ΔAIC_c	PIGG ΔAIC_c	GOOSE ΔAIC_c	OTTER ΔAIC_c	LSMITH ΔAIC_c	USMITH ΔAIC_c	NOTT ΔAIC_c
1	12912.8	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
2	6459.4	521.8	52.5	42.1	29.5	13.2	13.0	37.9
3	3786.6	1468.7	36.4	15.1	43.2	38.5	45.5	67.2
4	1975.3	687.0	203.2		197.4	18.5		82.1
5	969.9	794.1			474.2	22.0		14.5
6	252.2	2902.0			484.7	104.0		64.4
7	**0.0**	128.3				17.2		
8	281.3	4123.3				34.3		
9	503.7	3344.5				25.2		
10	911.8	48.1				26.2		
11	1353.3	3764.5						
12	1641.8	4244.8						
13	2685.2	5911.0						
14	2982.3	62.5						
15	2943.3	3778.3						

Table B1. Results of tests for isolation-by-distance among pairs of *Percina rex* populations, sites, and individuals. Response variables included linearized F_{ST} and individual differentiation ($\hat{a}$); predictor variables included waterway distance, presence or absence of an intervening dam, and the harmonic-mean contemporary effective population size (N_e) of the pair. Simple and partial Mantel test results are based on 10^4 random matrix permutations. Statistically significant associations ($P < 0.05$) are indicated in bold. Population codes follow Table 1.

Unit being compared	Number of comparisons	Response	Predictor	Mantel *r*
Populations	21	F_{STL}	Distance	0.355
Populations	21	F_{STL}	N_e	**-0.491**
Populations	21	F_{STL}	N_e | Distance	**-0.494**
Populations	21	F_{STL}	Distance | N_e	0.360
All sites	595	F_{STL}	Distance	**0.434**
All sites	595	F_{STL}	Dam	**0.606**
All sites	595	F_{STL}	Dam | Distance	**0.491**
All sites	595	F_{STL}	Distance | Dam	0.163
Sites within populations	122	F_{STL}	Distance	0.250
Sites within UROAN	91	F_{STL}	Distance	-0.182
Sites within PIGG	3	F_{STL}	Distance	0.570
Sites within OTTER	10	F_{STL}	Distance	-0.124
Sites within LSMITH	10	F_{STL}	Distance	0.140
Sites within NOTT	6	F_{STL}	Distance	0.916
Individuals within UROAN	21945	$\hat{a}$	Distance	0.004
Individuals within PIGG	4186	$\hat{a}$	Distance	0.045
Individuals within GOOSE	561	$\hat{a}$	Distance	0.163
Individuals within OTTER	4560	$\hat{a}$	Distance	0.037
Individuals within LSMITH	1081	$\hat{a}$	Distance	0.103
Individuals within USMITH	861	$\hat{a}$	Distance	-0.012
Individuals within NOTT	1596	$\hat{a}$	Distance	0.076

APPENDIX C: Results of pedigree reconstruction on simulated datasets

Table C1. Proportions (number correct / number possible) of correctly matched family pairs and correctly unmatched unrelated pairs in pedigree reconstruction analysis of simulated datasets. Dataset characteristics and modeling parameters are described in the text.

Dataset	True mating system	Assumed mating system	Parents included	Error rate	Full siblings	Half siblings	Father-offspring	Mother-offspring	Unrelated
1	Monogamous	Monogamous	Yes	0	1.00 (42/42)		1.00 (42/42)	1.00 (42/42)	1.00 (1911/1911)
1	Monogamous	Monogamous	Yes	0.05	1.00 (42/42)		1.00 (42/42)	1.00 (42/42)	1.00 (1911/1911)
1	Monogamous	Monogamous	No	0	1.00 (42/42)				1.00 (1911/1911)
1	Monogamous	Monogamous	No	0.05	1.00 (42/42)				1.00 (1911/1911)
1	Monogamous	Promiscuous	Yes	0	1.00 (42/42)		1.00 (42/42)	1.00 (42/42)	1.00 (1911/1911)
1	Monogamous	Promiscuous	Yes	0.05	1.00 (42/42)		1.00 (42/42)	1.00 (42/42)	1.00 (1911/1911)
1	Monogamous	Promiscuous	No	0	0.60 (25/42)				1.00 (1911/1911)
1	Monogamous	Promiscuous	No	0.05	0.64 (27/42)				1.00 (1911/1911)
2	Polygynous	Monogamous	Yes	0	1.00 (14/14)	0.00 (0/28)	0.14 (4/28)	1.00 (28/28)	1.00 (1064/1064)
2	Polygynous	Monogamous	Yes	0.05	1.00 (14/14)	0.00 (0/28)	0.14 (4/28)	1.00 (28/28)	1.00 (1064/1064)
2	Polygynous	Monogamous	No	0	0.79 (11/14)	0.00 (0/28)			1.00 (1064/1064)
2	Polygynous	Monogamous	No	0.05	0.71 (10/14)	0.00 (0/28)			1.00 (1064/1064)
2	Polygynous	Promiscuous	Yes	0	1.00 (14/14)	1.00 (28/28)	1.00 (28/28)	1.00 (28/28)	1.00 (1064/1064)
2	Polygynous	Promiscuous	Yes	0.05	1.00 (14/14)	1.00 (28/28)	1.00 (28/28)	1.00 (28/28)	1.00 (1064/1064)
2	Polygynous	Promiscuous	No	0	0.57 (8/14)	0.68 (19/28)			1.00 (1064/1064)
2	Polygynous	Promiscuous	No	0.05	0.64 (9/14)	0.71 (20/28)			1.00 (1064/1064)
3	Promiscuous	Monogamous	Yes	0	0.93 (26/28)	0.00 (0/112)	0.32 (18/56)	0.36 (20/56)	1.00 (2856/2856)
3	Promiscuous	Monogamous	Yes	0.05	1.00 (28/28)	0.00 (0/112)[a]	0.57 (32/56)	0.50 (28/56)	1.00 (2856/2856)
3	Promiscuous	Monogamous	No	0	0.82 (23/28)	0.00 (0/112)[b]			1.00 (2856/2856)
3	Promiscuous	Monogamous	No	0.05	0.86 (24/28)	0.00 (0/112)[c]			1.00 (2856/2856)

Table C1, continued

3	Promiscuous	Promiscuous	Yes	0	1.00 (28/28)	1.00 (112/112)	1.00 (56/56)	1.00 (56/56)	1.00 (2856/2856)
3	Promiscuous	Promiscuous	Yes	0.05	1.00 (28/28)	1.00 (112/112)	1.00 (56/56)	1.00 (56/56)	1.00 (2856/2856)
3	Promiscuous	Promiscuous	No	0	0.96 (27/28)	0.96 (108/112)			1.00 (2856/2856)
3	Promiscuous	Promiscuous	No	0.05	0.96 (27/28)	0.96 (108/112)			1.00 (2856/2856)
4	Unrelated	Monogamous	Yes	0			1.00 (14/14)	1.00 (14/14)	1.00 (91/91)
4	Unrelated	Monogamous	Yes	0.05			1.00 (14/14)	1.00 (14/14)	1.00 (91/91)
4	Unrelated	Monogamous	No	0					1.00 (91/91)
4	Unrelated	Monogamous	No	0.05					1.00 (91/91)
4	Unrelated	Promiscuous	Yes	0			1.00 (14/14)	1.00 (14/14)	1.00 (91/91)
4	Unrelated	Promiscuous	Yes	0.05			1.00 (14/14)	1.00 (14/14)	1.00 (91/91)
4	Unrelated	Promiscuous	No	0					1.00 (91/91)

[a]16 half-sib pairs erroneously were matched as full sibs

[b]3 half-sib pairs erroneously were matched as full-sibs

[c]28 half-sib pairs erroneously were matched as full-sibs

APPENDIX D: Summary of microsatellite genetic diversity statistics in *Percina rex*

Table D1. Statistics include expected (H_E) and observed (H_O) heterozygosity, richness of alleles (A), and the ratio of allele richness to allele size-range (M) estimated for each locus in each population. H_O values significantly lower than expected under Hardy-Weinberg equilibrium (based on 10,000 permutations and an alpha of 0.01) are shown in bold.

Statistic	Population	Prex33	Prex37	Prex45	Prex42	Prex46	Prex36	Prex38	Prex41	Prex43	Prex44	Prex47
H_E	UROAN	0.81	0.69	0.86	0.89	0.92	0.78	0.86	0.87	0.67	0.90	0.91
	PIGG	0.28	0.69	0.65	0.56	0.63	0.81	0.58	0.78	0.66	0.71	0.79
	GOOSE	0.13	0.64	0.61	0.60	0.19	0.71	0.74	0.69	0.31	0.62	0.63
	OTTER	0.65	0.65	0.64	0.44	0.43	0.30	0.65	0.74	0.33	0.58	0.71
	LSMITH	0.39	0.80	0.69	0.78	0.69	0.42	0.76	0.76	0.54	0.79	0.81
	USMITH	0.67	0.81	0.78	0.82	0.85	0.81	0.85	0.78	0.71	0.81	0.84
	NOTT	0.57	0.71	0.71	0.89	0.86	0.65	0.72	0.75	0.67	0.47	0.76
H_O	UROAN	0.90	0.67	0.90	0.87	0.80	0.70	0.83	0.87	0.57	**0.60**	0.87
	PIGG	0.30	0.73	0.73	0.43	0.72	0.70	0.67	**0.67**	0.57	0.67	0.83
	GOOSE	**0.00**	0.63	0.59	0.50	0.17	0.67	0.72	0.90	0.30	0.63	0.67
	OTTER	0.60	0.80	0.70	0.47	0.40	0.30	0.77	0.83	0.33	0.57	0.77
	LSMITH	0.33	0.87	0.67	0.60	0.57	0.40	0.77	0.83	0.43	0.70	0.70
	USMITH	0.66	0.77	0.80	0.87	0.87	0.90	0.87	0.83	0.67	0.70	**0.57**
	NOTT	0.55	0.70	0.67	0.87	0.83	0.73	0.77	0.83	0.60	0.43	0.70
A	UROAN	7	7	10	13	14	8	9	12	5	13	13
	PIGG	5	5	6	4	5	7	5	6	3	4	9
	GOOSE	2	3	4	4	3	5	4	5	3	4	5
	OTTER	3	4	3	3	2	2	3	4	2	3	4
	LSMITH	4	5	6	6	6	4	5	5	5	5	7
	USMITH	4	9	9	8	11	8	8	8	5	11	9
	NOTT	5	5	5	11	10	3	5	7	6	3	8
M	UROAN	0.70	0.88	1.00	0.72	0.82	1.00	0.64	1.00	1.00	0.72	1.00
	PIGG	0.83	0.63	0.86	0.40	0.56	0.88	0.45	0.67	1.00	0.44	0.64
	GOOSE	0.67	0.43	0.57	0.40	0.50	0.63	0.36	0.56	1.00	0.44	0.45
	OTTER	1.00	0.57	1.00	0.33	0.33	1.00	0.27	0.57	0.67	0.38	0.50
	LSMITH	1.00	1.00	0.75	0.60	0.33	0.33	0.36	0.56	0.71	0.45	0.58
	USMITH	0.67	0.69	0.82	0.67	0.46	0.89	0.73	0.57	0.45	0.61	0.64
	NOTT	1.00	0.45	0.56	0.69	0.91	0.50	0.83	0.88	1.00	0.33	1.00

APPENDIX E: Summary of *ND2* mitochondrial DNA study of *Percina rex*

METHODS

Sample collection

Field collections of Roanoke logperch (*Percina rex*) were made by me, the Virginia Department of Game and Inland Fisheries, and the North Carolina Wildlife Resources Commission between 2003 and 2008 in all portions of the species' known range. Fish were captured by backpack or barge electrofisher and either a seine or dipnet, using methods approved by the U.S. Fish and Wildlife Service (USFWS). We temporarily anesthetized all captured fish in MS-222 (Finquel), removed a 5-mm x 5-mm section of tissue from the dorsal half of the caudal fin, and then returned fish alive to the locality of capture. Tissue samples were dried in coin envelopes and then stored at -20ºC until DNA extraction. Template DNA subsequently was extracted from whole tissue samples using a PureGene DNA Extraction Core Kit A (Gentra Systems, Minneapolis, Minnesota, USA), according to the manufacturer's instructions. A total of 578 individual DNA samples were collected from fish at a total of 35 spatial localities and analyzed to estimate the population genetic structure of the species based on microsatellite genetic markers (see Chapter 1). For mtDNA characterization, I used a stratified random subsample of 30 individuals from each of the seven populations of Roanoke logperch inferred in Chapter 1, for a total of 210 fish analyzed. Populations were stratified by spatial sub-regions (e.g., sub-watersheds, streams), and then a roughly equal number of individuals were randomly selected from each component sub-region.

Laboratory analyses

I directly sequenced the light and heavy strands of the 1047-bp *ND2* mtDNA gene. Forward and reverse primers for PCR were *ND2* 562L and *ND2* 449H, respectively, from George et al. (2006). PCR employed 25-µl reactions with the following reagent mix: 2 µL of 2.5-mM each dNTPs (premixed); 2.5 µL of 10X NH$_4$ Ex*Taq* buffer (MgCl$_2$ included); 1 µL each of 20-µM ND2 562L and ND2 449H primers; 0.15 µL of 5 Units µL^{-1} Ex*Taq* polymerase (TaKaRa Bio, Inc., Otsu, Shiga, Japan); 3 µL of 20-ng µL^{-1} template DNA; and 15.35 µL of dH$_2$O. PCR was conducted in a MyCycler Thermal Cycler (BioRad, Hercules, California, USA) by using an initial denaturation step (94°C, 3 min), followed by 35 cycles of

denaturation (94°C, 40 sec), annealing (60°C, 40 sec), and extension (72°C, 60 sec), and a final extension step (72°C, 2 min). Non-specific amplification products were removed with ExoSAP-IT (USB Corp., Cleveland, Ohio, USA) and the cleaned DNA was diluted to 10 ng μL^{-1} for forward and reverse sequencing in an ABI 3130 automated sequencer (Applied Biosystems, Inc., Foster City, California, USA).

Data analysis

Forward and reverse sequence fragments were aligned and edited from the raw electropherograms using SEQUENCHER version 3.0 (Gene Codes Corp., Ann Arbor, Michigan, USA). Only the central 1037 bp (i.e., positions 6-1042) of aligned sequences could be reliably interpreted, so I retained only this region for analyses. Haplotypes were deposited in the GenBank public database (accession numbers JF929000-JF929014).

Haplotype (Nei and Tajima 1981) and nucleotide (Nei 1987) diversity were estimated for each population and overall using DNAsp version 5.1 (Librado and Rozas 2009). Genetic distances between *ND2* haplotypes were estimated with the *p*-distance method in MEGA 5.05 (http://www.megasoftware.net/). I used two tests to examine whether *ND2* variation within Roanoke logperch was consistent with selective neutrality: 1) Fisher's exact test of the null hypothesis that nonsynonymous mutations $\leq$ synonymous mutations between each pair of haplotypes in MEGA, using the Nei-Gojobori method, and 2) a two-tailed test of the null hypothesis that Tajima's $D = 0$ using DNAsp, assuming a beta distribution.

Evolutionary relationships among haplotypes were inferred in a maximum parsimony framework. Maximum parsimony methods assume no particular nucleotide mutation model, but determine the simplest evolutionary history that is consistent with the observed data. I used TCS version 1.21 (Clement et al. 2000) to construct a haplotype network based on maximum parsimony criteria and a confidence limit of 95%.

RESULTS

Among the 210 individuals analyzed, 15 different haplotypes were observed (Table E1). Of the 1037 nucleotide sites in the sequence, 22 of these sites (2%) were variable. A total of 23 mutations were observed (nucleotide position 662 exhibited three states), of which 6 were nonsynonymous. The ratio of nonsynonymous to synonymous mutations was 0.35, which is consistent with a lack of positive selection on the *ND2* haplotypes (Ford 2002). Furthermore, the overall Tajima's *D* value of -0.42 did not deviate significantly from zero ($P > 0.1$), which is consistent with selective neutrality. I therefore interpreted observed *ND2* variation as selectively neutral, with patterns of variation resulting from demographic processes.

The overall haplotype (*h*) and nucleotide (π) diversity of 210 *Percina rex* individuals at *ND2* was 0.775 and 0.0032, respectively (Table E2). Genetic diversity varied widely among populations. The upper Roanoke River population (UROAN) exhibited by far the greatest haplotype and nucleotide diversity, whereas the Otter River population (OTTER) exhibited no genetic variation (Table E2). Eleven of the fifteen haplotypes observed were found in UROAN, and eight of these were unique to UROAN. Only two other populations (LSMITH and NOTT) exhibited unique haplotypes.

Based on maximum parsimony criteria, all haplotypes exhibited relatively close evolutionary relationships (Figure E1). Haplotypes were separated by one to nine mutation events (0.1-0.9% divergence). There appeared to be five primary clades, with little geographic population structure to haplotype relationships. UROAN individuals were represented in three of the five clades. UROAN shared three haplotypes with the geographically distant USMITH, and shared at least one haplotype with all populations except OTTER and NOTT. OTTER, on the other hand, exhibited only one haplotype, which was shared with PIGG and GOOSE. The only population that clustered separately from all other populations was NOTT, which also is the most geographically disjunct of the seven populations. However, the NOTT clade was only 0.3-0.8% divergent from other clades.

REFERENCES

Clement, M., Posada, D., and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1660.

Ford, M. J. 2002. Applications of selective neutrality tests to molecular ecology. Molecular Ecology 11:1245-1262.

Jenkins, R. E., and N. M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Librado, P. and J. Rozas. 2009. DNAsp version 5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451-1452.

Moser, G. A. 1992. Roanoke logperch (*Percina rex*) recovery plan. Final Report to the U.S. Fish and Wildlife Service, Newton Corner, Massachusetts.

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

Nei, M., and F. Tajima. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25:1253-1256.

Roberts, J. H., Angermeier, P. L., and E. M. Hallerman. 2009. Analysis of population genetics of Roanoke logperch (*Percina rex*) based on microsatellite markers. Final Report to U.S. Fish and Wildlife Service, Gloucester, Virginia.

Table E1. Summary of nucleotide variation at the sequenced 1037-bp region of the *ND2* mtDNA gene in *Percina rex*. Numbers of individuals (*n*) bearing each haplotype (lettered A-O) are indicated. Table entries show nucleotide substitutions relative to the most-commonly occurring haplotype (Haplotype A), whether the substitution was a transition (*s*) or transversion (*v*), and the position of the site within a codon. Nonsynonymous substitutions are shown in bold typeface. Positions of variable nucleotide sites should be read vertically (i.e., 48, 84, 292, etc.).

		Variable nucleotide sites																						
				2	2	4	4	5	6	6	7	7	7	8	8	8	8	8	9	9	9	9	9	
		4	8	9	9	7	8	6	2	6	1	2	8	0	2	6	7	7	3	3	5	9	9	GenBank
Haplotype	*n*	8	4	2	7	5	9	4	1	2	7	3	6	4	2	1	3	6	3	9	0	0	2	accession #
A	75	C	C	A	T	G	C	G	T	G	A	C	A	A	C	T	T	T	G	C	G	A	C	JF929001
B	4	.	.	**G**	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	JF929005
C	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	JF929011
D	1	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	**T**	JF929013
E	1	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	JF929010
F	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	JF929004
G	3	.	.	.	.	.	.	.	.	**C**	.	.	.	.	.	.	.	.	.	.	.	.	.	JF929014
H	1	.	.	.	.	.	.	.	.	.	.	A	.	G	.	.	C	.	.	.	.	.	.	JF929003
I	60	.	.	.	.	**A**	A	.	.	.	G	.	.	G	.	.	.	.	.	.	.	.	.	JF929000
J	16	.	.	.	.	**A**	.	.	.	.	G	.	.	G	.	.	.	.	.	.	.	G	.	JF929002
K	18	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	C	.	.	**A**	.	.	JF929006
L	12	.	.	.	.	.	.	A	.	.	.	.	.	G	.	.	.	C	.	.	**A**	.	.	JF929007
M	10	.	.	.	.	.	.	A	C	.	.	.	.	G	A	.	.	.	.	.	.	.	.	JF929009
N	2	.	T	.	C	.	.	A	C	.	.	.	.	G	A	C	.	.	.	.	.	.	.	JF929012
O	4	.	.	.	.	.	.	A	C	**A**	.	.	.	G	A	.	.	.	.	.	.	.	.	JF929008
Substitution		*v*	*s*	*s*	*s*	*s*	*v*	*s*	*s*	*s/v*	*s*	*v*	*s*	*s*	*v*	*s*	*s*	*s*	*s*	*s*	*s*	*s*	*s*	
Position		3	3	1	3	1	3	3	3	2	3	3	3	3	3	3	3	3	3	3	2	3	2	

Table E2. Summary of genetic diversity statistics for seven populations of *Percina rex* at the *ND2* mtDNA gene. Entries include the sample size of individuals (n) and the observed number of *ND2* haplotypes (K), segregating sites (S), haplotype diversity (H_d), and nucleotide diversity (π).

Population	n	K	Private haplotypes	S	H_d	π
UROAN	30	11	8	16	0.885	0.0031
PIGG	30	2	0	4	0.460	0.0018
GOOSE	30	2	0	4	0.331	0.0013
OTTER	30	1	0	0	0.000	0.0000
LSMITH	30	2	1	4	0.515	0.0020
USMITH	30	3	0	5	0.393	0.0016
NOTT	30	2	2	1	0.497	0.0005
Total	210	15		22	0.775	0.0032
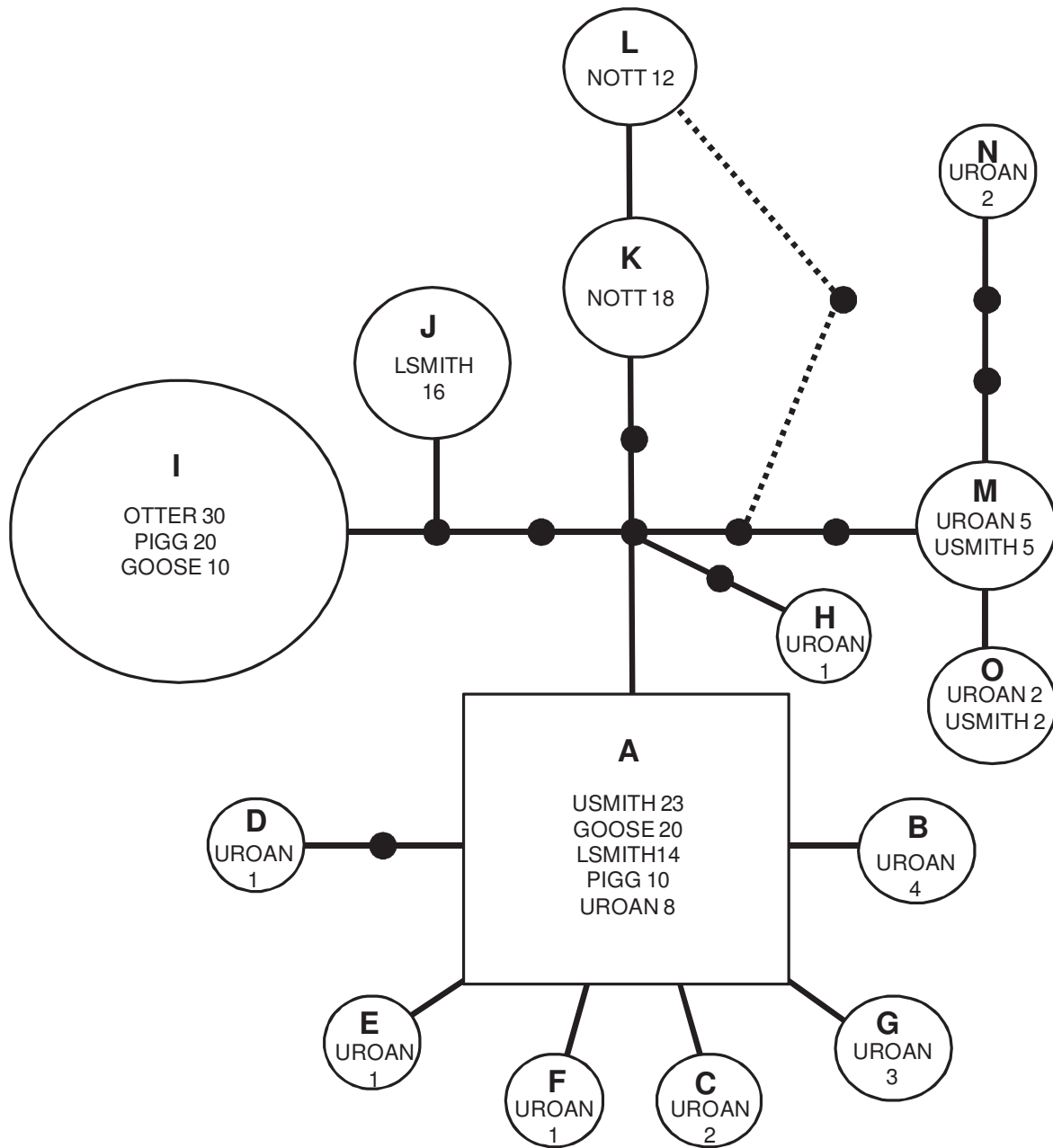
Figure E1. Maximum parsimony haplotype network for the sequenced 1037-bp region of the *ND2* mtDNA gene in *Percina rex*. Polygons indicate observed haplotypes, line segments indicate hypothesized individual mutation events between haplotypes, and nodes indicate hypothesized unobserved haplotypes. The dotted line segment indicates a parsimonious alternative mutation pathway. Haplotype names (capital letters in bold) follow those of Table E1. Numerals indicate the number of individuals (out of 30) from each population that bore a given haplotype. Population codes are explained in the text.

APPENDIX F: Summary of approximate Bayesian computation models for *Percina rex*

APPROACH

I used approximate Bayesian computation (ABC) models to estimate demographic parameters of interest for assessing risk to *Percina rex* (Beaumont et al. 2002). The principle of ABC is to iteratively simulate millions of demographic scenarios, each iteration drawing parameter estimates (e.g., N_e, m) from plausible prior distributions. From each simulated dataset, a series of genetic summary statistics are calculated. Following all simulation runs, these simulated statistics are compared to empirical statistics calculated from real populations, and simulations providing a poor match are rejected. Posterior estimates of demographic parameters are then obtained from the subset of simulations that has been retained.

I parameterized demographic simulations based on the presumed history of *P. rex*. There presently are seven populations that are organized into three major basins (Roanoke, Dan, and Nottoway). Each population was assigned a uniform prior distribution on effective population size (N_e) between 2 and 5000 for microsatellite simulations and a uniform prior distribution on female N_e between 1 and 2500 for mtDNA simulations.

These seven populations are completely isolated by dams, reservoirs, and/or unsuitable habitat (i.e., the contemporary migration rate [m] is zero; see Chapter 1), but may have been connected by migration historically. I assumed that these historical migration rates were symmetrical within population-pairs but could vary between population-pairs depending on whether the comparison was 1) within a basin, 2) between the Roanoke and Dan basins, 3) between the Roanoke and Nottoway basins, or 4) between the Dan and Nottoway basins. In each case, I assigned a uniform prior to historical m between 0 and 0.05.

The transition from current isolation to historical migration occurred at some time t (measured in generations) looking backward in time from the present. I assigned a uniform prior between 10 and 80 generations to t. Given that *Percina rex* matures at 2.5 years and lives to 6.5 years, I assumed that its generation time was the midpoint of this interval, 4.5 years. The prior distribution of t thus corresponded to a range of 45 to 360 before present. Given that most samples were collected in 2005, this range of

dates (1960-1645) puts the fragmentation event somewhere between the time that most Virginia and North Carolina reservoirs were constructed (i.e., 1920-1964) and prior the onset of alterations of the environment by European settlers (i.e., 1700s; Jenkins and Burkhead 1994).

In addition to fragmentation, I assumed that populations may have undergone a bottleneck at time t. I allowed this single-generation reduction (B) to range from 0 to 90% of historical N_e. Because simulations ran backwards in time, this was accomplished by applying a multiplier of between 1 and 9 to each population's contemporary N_e.

In addition to these 19 demographic parameters, I also needed to specify a mutation rate and model for each marker type. For microsatellites, I assumed a stepwise mutation model and a uniform prior distribution of mutation rate between 0.0001 and 0.01 (Ellegren 2004). For mtDNA, I estimated the appropriate mutation model for *ND2* in *P. rex* using MEGA 5.05 (http://www.megasoftware.net/), and found it to be a general time -reversible model with rate variation among sites following a gamma distribution of shape parameter 0.05. I assigned a uniform prior distribution of mtDNA mutation rate between 1.3×10^{-7} and 3.1×10^{-5} per site (Denver et al. 2000; Lambert et al. 2002; Howell et al. 2003), which corresponded to a range of 0.00013 to 0.0321 across the whole locus (i.e., 1037 sites).

Simulated datasets were generated in BayeSSC (http://www.stanford.edu/group/hadlylab/ssc/index.html) by randomly drawing values from these prior distributions and simulating demographies backwards in time to coalescence. During the ABC rejection step, summary statistics resulting from simulations were compared to empirical statistics calculated from *P. rex* populations (see Chapter 3). The Euclidean distance between each empirical and simulated statistic was calculated and then summed across statistics within a given simulation iteration. For microsatellites, rejection was based on allele richness, gene diversity, and R_{ST} (Slatkin 1995). For mtDNA, rejection was based on haplotype and nucleotide diversity, number of segregating sites, number of private haplotypes, number of pairwise differences, and F_{ST} (Hudson et al. 1992). The rejection procedure was performed in R 2.10.0 (R Core Development Team) using a script written by C. Anderson (Harvard University) and a tolerance threshold of 0.05%. I simulated a total of 5,000,000 demographies

for each marker type, updating priors after every 1,000,000 simulations. Posterior distributions of

demographic parameters were estimated from the final 500 accepted simulations. A summary of prior

and posterior distributions is presented in Table F1. Densities of prior and posterior parameter estimates

are shown in Figures F1 and F2.

REFERENCES

Beaumont, M.A., W. Zhang, and D.J. Balding. 2002. Approximate Bayesian computation in population

genetics. Genetics 162:2025-2035.

Denver, D.R., K. Morris, M. Lynch, L.L. Vassilieva, and W.K. Thomas. 2000. High direct estimate of

the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. Science 289:2342-

2344.

Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics

5:435-445.

Howell, N., C.B. Smejkal, D.A. Mackey, P.F. Chinnery, D.M. Turnbull, and C. Herrnstadt. 2003. The

pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference

between phylogenetic and pedigree rates. American Journal of Human Genetics 72:659-670.

Hudson, R.R., M. Slatkin, and W.P. Maddison. 1992. Estimation of levels of gene flow from DNA

sequence data. Genetics 132:583-589.

Jenkins, R.E., and N.M. Burkhead. 1994. Freshwater Fishes of Virginia. American Fisheries Society,

Bethesda, MD.

Lambert, D.M., P.A. Ritchie, C.D. Millar, B. Holland, A.J. Drummond, and C. Baroni. 2002. Rates of

evolution in ancient DNA from Adelie penguins. Science 295:2270-2273.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies.

Genetics 139:457-462.

Table F1. Characteristics of uniform prior and estimated posterior distributions for parameters used in approximate Bayesian computation demographic simulations for *Percina rex*. Parameters include the contemporary effective populations size (N_e), bottleneck severity (B), historical migration rate (m), and number of generations in the past in which fragmentation and bottlenecks occurred (t).

Parameter	Microsatellites		mtDNA	
		Posterior		Posterior
	Prior (range)	mode (95% interval)	Prior (range)	mode (95% interval)
UROAN N_e	U:(2, 5000)	1198 (365, 3143)	U:(1, 2500)	280 (140, 1042)
PIGG N_e	U:(2, 5000)	601 (202, 2285)	U:(1, 2500)	72 (25, 338)
GOOSE N_e	U:(2, 5000)	11 (2, 144)	U:(1, 2500)	20 (8, 73)
OTTER N_e	U:(2, 5000)	99 (22, 428)	U:(1, 2500)	5 (1, 36)
LSMITH N_e	U:(2, 5000)	698 (237, 2478)	U:(1, 2500)	51 (18, 290)
USMITH N_e	U:(2, 5000)	196 (39, 621)	U:(1, 2500)	38 (16, 109)
NOTT N_e	U:(2, 5000)	300 (61, 1104)	U:(1, 2500)	11 (1, 59)
UROAN B	U:(0, 0.9)	0.51 (0.03, 0.88)	U:(0, 0.9)	0.47 (0.02, 0.87)
PIGG B	U:(0, 0.9)	0.26 (0.01, 0.87)	U:(0, 0.9)	0.49 (0.03, 0.88)
GOOSE B	U:(0, 0.9)	0.13 (0.02, 0.86)	U:(0, 0.9)	0.66 (0.03, 0.88)
OTTER B	U:(0, 0.9)	0.33 (0.03, 0.88)	U:(0, 0.9)	0.60 (0.03, 0.88)
LSMITH B	U:(0, 0.9)	0.22 (0.02, 0.88)	U:(0, 0.9)	0.36 (0.03, 0.88)
USMITH B	U:(0, 0.9)	0.73 (0.02, 0.87)	U:(0, 0.9)	0.23 (0.04, 0.88)
NOTT B	U:(0, 0.9)	0.21 (0.02, 0.87)	U:(0, 0.9)	0.74 (0.02, 0.88)
Within-basin m	U:(0, 0.05)	0.021 (0.001, 0.049)	U:(0, 0.05)	0.041 (0.002, 0.049)
Roanoke-Dan m	U:(0, 0.05)	0.006 (0.001, 0.049)	U:(0, 0.05)	0.025 (0.001, 0.048)
Roanoke-Nottoway m	U:(0, 0.05)	0.039 (0.002, 0.049)	U:(0, 0.05)	0.012 (0.001, 0.049)
Dan-Nottoway m	U:(0, 0.05)	0.013 (0.002, 0.049)	U:(0, 0.05)	0.022 (0.001, 0.049)
Generation of event t	U:(10, 80)	14 (10, 67)	U:(10, 80)	17 (10, 37)
Mutation rate per locus	U:(0.0001,0.01)	0.0028 (0.0007, 0.0104)	U:(0.00013,0.0321)	0.001 (0.0006, 0.0047)
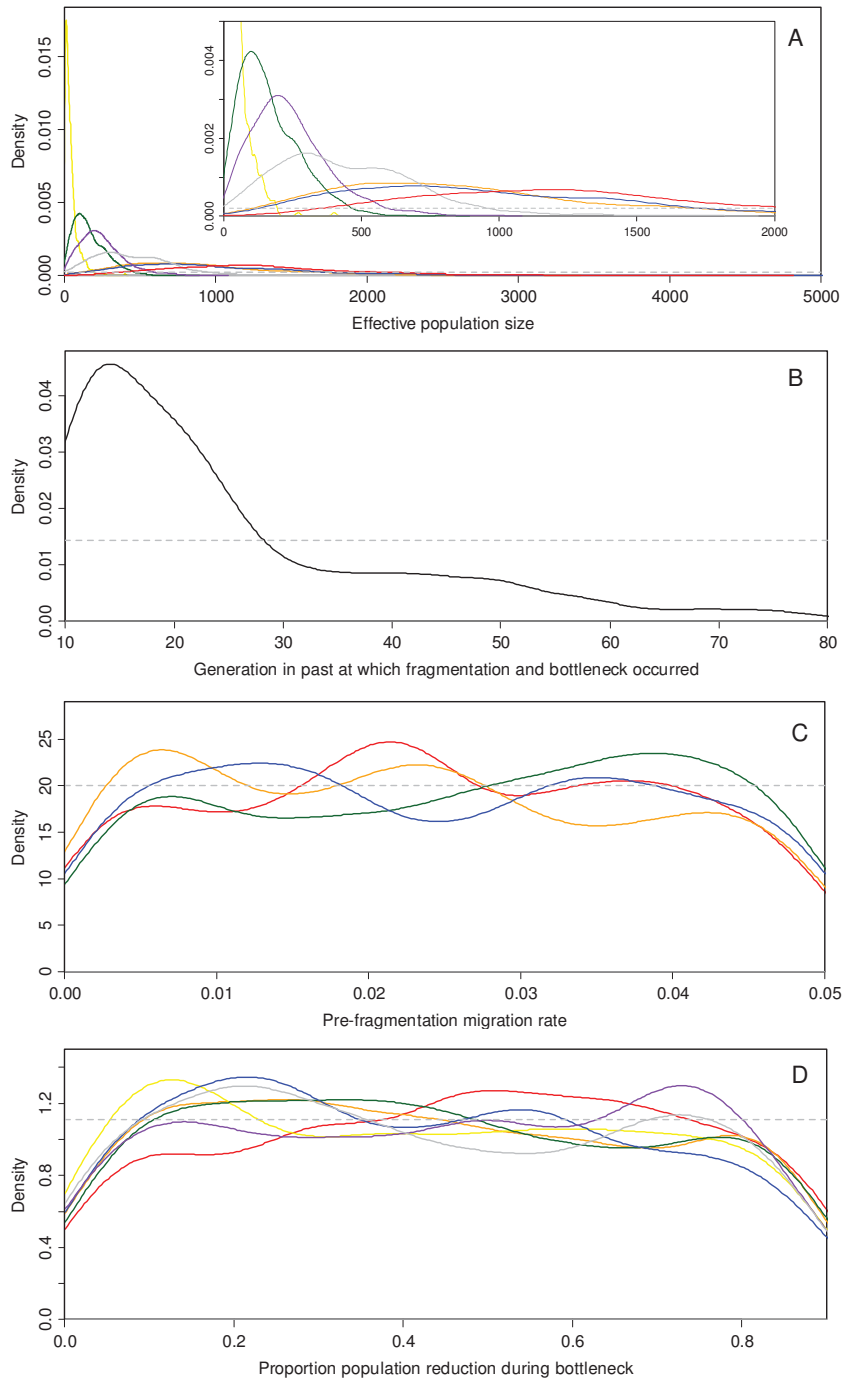
Figure F1. Prior (gray dotted lines) and posterior (solid lines) parameter densities (i.e., observed frequencies) from approximate Bayesian computation simulations of *Percina rex* demographic history based on microsatellite data. The inset in panel A shows a reduced axis range, for clarity. Color-coding schemes for panels A and D are as follows: UROAN (red), PIGG (orange), GOOSE (yellow), OTTER (green), LSMITH (blue), USMITH (purple), and NOTT (gray). Color-coding for panel C is as follows: within basins (red), between Roanoke and Dan (orange), between Roanoke and Nottoway (green), and between Dan and Nottoway (blue).
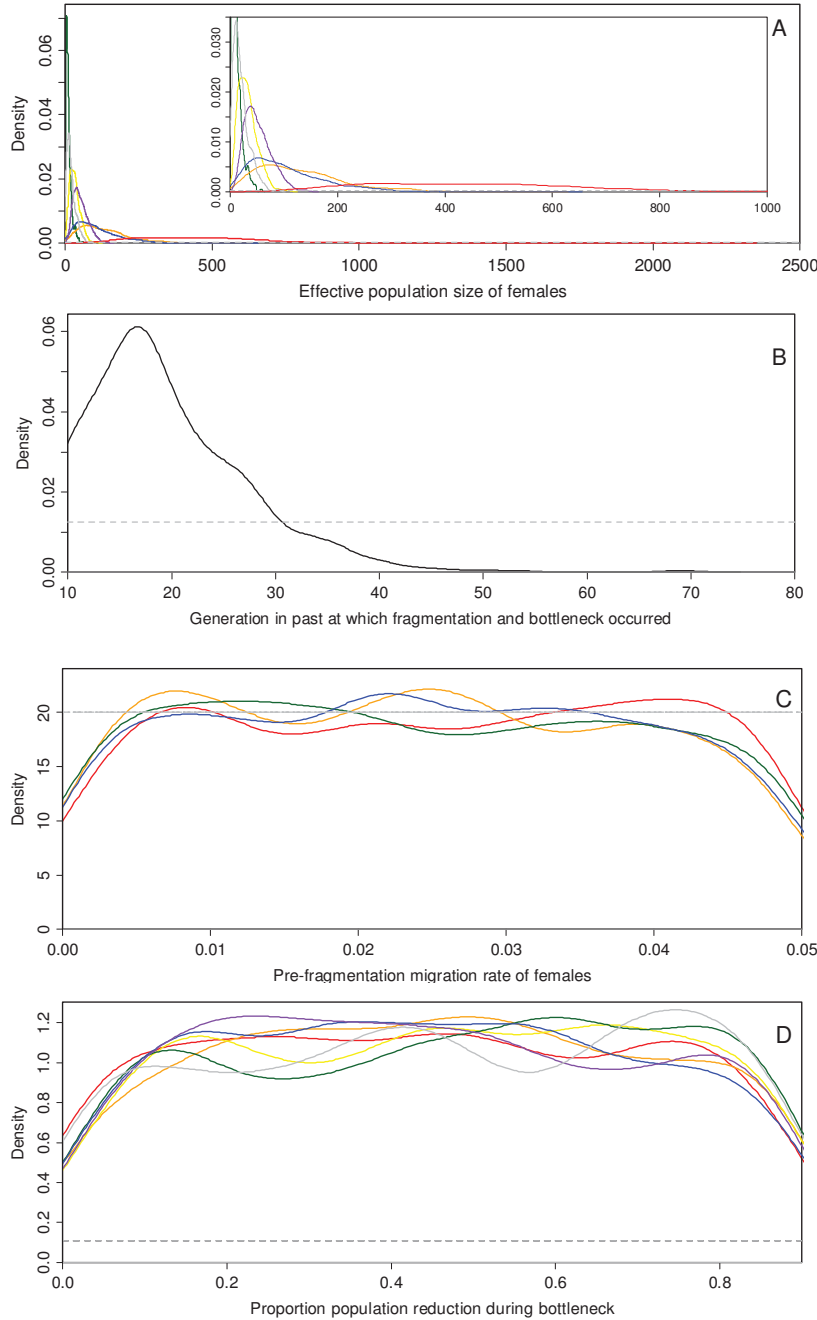
Figure F2. Prior (gray dotted lines) and posterior (solid lines) parameter densities (i.e., observed frequencies) from approximate Bayesian computation simulations of *Percina rex* demographic history based on mtDNA data. The inset in panel A shows a reduced axis range, for clarity. Color-coding schemes for panels A and D are as follows: UROAN (red), PIGG (orange), GOOSE (yellow), OTTER (green), LSMITH (blue), USMITH (purple), and NOTT (gray). Color-coding for panel C is as follows: within basins (red), between Roanoke and Dan (orange), between Roanoke and Nottoway (green), and between Dan and Nottoway (blue).

170

APPENDIX G: Summary of estimation of total population size for *Percina rex*

APPROACH

No estimates of absolute population size (*N*) have previously been developed for any *Percina rex* population. Such estimates would be useful for assessing small-population risks due to low *N*. I derived coarse estimates of *N* based on published and unpublished estimates of range extent, habitat availability, and *P. rex* relative abundance in populations. These estimates are based on best available scientific data, but are provisional and should be supplanted by better estimates when and if such data become available.

First, I estimated the total geographic extent of each population of *P. rex* based on distributional limits given by Rosenberger (2007) and personal communication with agency personnel (see Roberts et al. 2009). Once the upstream and downstream distributional limits were determined for a population, I calculated range extent as the total length of stream (km) between these points, measured in Google Earth 5.1 (http://www.google.com/earth) (Table G1).

P. rex primarily occupies pool habitat patches in the NOTT population and riffle-run habitat patches elsewhere (Rosenberger 2002). I estimated the average catch (*c*) of adult *P. rex* per habitat patch (riffle-run or pool) for each population (Table G1). For five populations, *c* was estimated based on data collected via a standardized quadrat-based backpack electrofishing method conducted in riffle-runs. In this way, the average *c* per riffle-run was estimated for UROAN (139 collections over 2006-2011; Roberts and Angermeier, unpublished data), USMITH (30 collections over 2006-2011; Roberts and Angermeier, unpublished data), PIGG (25 collections over 2003-2005; Lahey and Angermeier 2007), GOOSE (25 collections over 2003-2005; Lahey and Angermeier 2007), and OTTER (15 collections over 2003-2005; Lahey and Angermeier 2007). NOTT has not been sampled using the quadrat-based method, but based on snorkeling, Rosenberger (2002) found *P. rex* to be ~ 45% less abundant per sampled patch in NOTT than in UROAN. I therefore assumed that *c* for NOTT was 0.55 times the *c* for UROAN. Data on *c* for LSMITH were unavailable, so I assumed that *c* for LSMITH was the same as *c* for the nearby USMITH population. Based on mark-recapture data, Roberts and Angermeier (unpublished data)

estimate the sampling efficiency of the quadrat-based method to be approximately 10%, so I multiplied all *c* estimates by 10 to estimate the total abundance of fish per patch.

Data on habitat patch density (riffle-runs km^{-1} or pools km^{-1}) were extracted from the sources above (Table G1). The patch density of LSMITH was assumed to be the same as for USMITH. I then multiplied patch density by the total abundance of fish per patch to estimate total logperch density (fish km^{-1}). Finally, I multiplied this density by range extent to estimate the total population size of adult logperch for each population (Table G1).

REFERENCES

Lahey, A.M., and P.L. Angermeier. 2007. Range-wide assessment of habitat suitability for Roanoke logperch (*Percina rex*). Final Report to Virginia Transportation Research Council, Charlottesville Virginia.

Roberts, J.H. 2009. Analysis of population genetics of Roanoke logperch (*Percina rex*) based on microsatellite markers. Final Report to U.S. Fish and Wildlife Service, Gloucester, VA.

Rosenberger, A.E. 2002. Multi-scale habitat use patterns of Roanoke logperch (*Percina rex*) in Virginia rivers: a comparison among populations and over ontogeny. Dissertation submitted to the Department of Fisheries and Wildlife Sciences, Virginia Tech, Blacksburg, VA.

Table G1. Estimated and derived demographic and habitat parameters for *Percina rex*. See text for details. Where applicable and available, I include both the mean and standard deviation (in parentheses) of parameters.

Population	Range extent (km)	Patch density (patches km^{-1})	Raw catch (fish patch^{-1})	Total fish density (fish km^{-1})	Total population size (fish)
UROAN	118	2.33 (0.36)	5.05 (3.57)	117.7	13884
PIGG	100	12.25 (4.09)	0.56 (1.04)	68.6	6860
GOOSE	40	10.10 (2.68)	0.40 (0.58)	40.4	1616
OTTER	53	9.07 (7.31)	0.33 (0.42)	29.9	1586
LSMITH	105	5.47	1.63	89.2	9362
USMITH	28	5.47 (0.49)	1.63 (1.85)	89.2	2497
NOTT	193	3.11 (2.99)	2.78	86.5	16686

APPENDIX H: Summary of microsatellite statistics in *Etheostoma rufilineatum and E. blennioides*

Table H1. Genetic diversity statistics for *E. rufilineatum.* Statistics include expected (H_E) and observed (H_O) heterozygosity, richness of alleles (A), and the ratio of allele richness to allele size-range (M) estimated for each locus in each population. H_O values significantly lower than expected under Hardy-Weinberg equilibrium (based on 10,000 permutations and an alpha of 0.01) are shown in bold. PCR annealing temperature and observed allele size-range (in basepairs) are given for each locus.

	Locus	*Esc26*	*Esc18*	*Esc132*	*E06*	*CV12*
	Annealing temp	56°	56°	56°	56°	56°
	Size range (bp)	152-260	86-182	142-234	122-174	147-235
Statistic						
H_E	BIGC	0.884	0.457	0.847	0.793	0.828
	BMOC	0.872	0.529	0.872	0.568	0.814
	BSYC	0.800	0.775	0.794	0.662	0.776
	BULL	0.901	0.787	0.877	0.736	0.870
	CLI1	0.922	0.790	0.870	0.717	0.918
	CLI2	0.905	0.801	0.870	0.788	0.861
	COPP	0.893	0.602	0.893	0.753	0.916
	EMOR	0.857	0.722	0.925	0.686	0.927
	FLAT	0.812	0.561	0.835	0.588	0.900
	LAUR	0.881	0.820	0.884	0.714	0.694
	LICK	0.903	0.668	0.871	0.826	0.822
	LITT	0.820	0.771	0.871	0.714	0.803
	LPIG	0.912	0.633	0.909	0.867	0.865
	MFHR	0.907	0.775	0.860	0.820	0.847
	NFH1	0.863	0.725	0.874	0.748	0.821
	NFH2	0.857	0.746	0.877	0.768	0.737
	POWE	0.891	0.760	0.899	0.791	0.907
	SFHR	0.892	0.813	0.897	0.753	0.863
	TEL1	0.765	0.570	0.872	0.637	0.830
	TEL2	0.836	0.663	0.845	0.507	0.745
	WACK	0.882	0.683	0.860	0.773	0.871
	WHIT	0.872	0.760	0.893	0.720	0.884
H_O	BIGC	0.792	0.458	0.875	0.875	0.750
	BMOC	0.826	0.522	0.913	0.652	0.739
	BSYC	0.739	0.870	0.783	0.739	0.826
	BULL	0.875	0.739	0.833	0.667	0.917
	CLI1	0.870	0.750	0.870	0.683	0.917
	CLI2	0.826	0.833	0.958	0.826	0.783
	COPP	0.913	0.565	0.913	0.739	0.913
	EMOR	0.917	0.708	0.875	0.667	1.000

	FLAT	0.739	0.545	0.870	0.565	0.883
	LAUR	0.958	0.796	0.875	0.750	0.750
	LICK	0.917	0.625	0.792	0.875	0.738
	LITT	0.857	0.824	0.833	0.739	0.762
	LPIG	0.860	0.700	0.917	0.917	0.952
	MFHR	0.826	0.826	0.917	0.870	0.870
	NFH1	0.833	0.782	0.917	0.667	0.833
	NFH2	0.917	0.667	0.917	0.750	0.803
	POWE	0.917	0.750	0.917	0.792	0.833
	SFHR	0.957	0.792	0.958	0.739	0.875
	TEL1	**0.500**	0.542	0.790	0.583	0.913
	TEL2	0.750	0.700	0.928	0.458	0.820
	WACK	0.870	0.591	0.833	0.739	0.917
	WHIT	0.875	0.750	0.870	0.683	0.875
Alleles	BIGC	10	6	11	8	9
	BMOC	12	3	11	5	9
	BSYC	10	6	11	5	9
	BULL	12	5	12	6	11
	CLI1	14	7	11	5	15
	CLI2	15	5	11	7	10
	COPP	10	6	13	8	11
	EMOR	11	6	12	6	17
	FLAT	7	3	10	4	10
	LAUR	11	8	9	4	6
	LICK	12	6	9	7	11
	LITT	10	8	11	6	12
	LPIG	13	5	12	10	11
	MFHR	13	8	12	8	8
	NFH1	11	6	10	8	12
	NFH2	10	6	9	7	7
	POWE	13	6	11	7	11
	SFHR	12	10	11	9	10
	TEL1	11	7	12	5	10
	TEL2	12	6	10	5	7
	WACK	13	6	9	7	10
	WHIT	9	6	13	6	10
M	BIGC	0.909	0.375	0.917	0.889	0.750
	BMOC	0.545	0.188	0.579	1.000	0.692
	BSYC	0.625	0.462	0.917	0.833	0.818

Table H1, continued

BULL	0.723	0.385	0.710	0.429	0.757
CLI1	0.933	0.350	0.917	0.833	0.750
CLI2	0.652	0.385	0.917	0.875	0.769
COPP	0.667	0.462	0.929	0.667	0.846
EMOR	0.786	0.462	0.600	0.545	0.810
FLAT	0.636	0.375	0.714	0.900	0.909
LAUR	0.917	0.500	0.900	1.000	0.667
LICK	0.923	0.462	0.750	0.333	0.550
LITT	0.769	0.400	0.846	0.600	0.923
LPIG	0.813	0.500	0.857	0.476	0.688
MFHR	0.867	0.500	0.857	0.421	0.727
NFH1	0.733	0.333	0.909	0.727	0.857
NFH2	0.769	0.462	0.818	1.000	0.583
POWE	0.650	0.462	1.000	1.000	0.917
SFHR	0.857	0.476	1.000	0.474	0.833
TEL1	0.733	0.538	0.923	1.000	0.625
TEL2	0.918	0.462	0.802	1.000	0.702
WACK	0.591	0.462	1.000	1.000	0.909
WHIT	0.818	0.464	0.834	0.366	0.809

Table H2. Genetic diversity statistics for *E. blennioides*. Statistics include expected (H_E) and observed (H_O) heterozygosity, richness of alleles (*A*), and the ratio of allele richness to allele size-range (*M*) estimated for each locus in each population. H_O values significantly lower than expected under Hardy-Weinberg equilibrium (based on 10,000 permutations and an alpha of 0.01) are shown in bold. PCR annealing temperature and observed allele size-range (in basepairs) are given for each locus.

	Locus	*Esc26*	*E09*	*CV09*	*Esc132*	*CV24*	*CV12*	*E06*
	Annealing temp	61°	61°	61°	61°	57°	57°	61°
	Size range (bp)	137-273	272-280	129-167	191-327	115-121	180-256	154-182
Statistic								
H_E	CLI1	0.918	0.340	0.813	0.932	0.444	0.913	0.859
	CLI2	0.932	0.428	0.726	0.948	0.428	0.957	0.806
	EMOR	0.873	0.736	0.795	0.946	0.191	0.941	0.630
	LITR	0.946	0.443	0.814	0.946	0.402	0.807	0.679
	LPIG	0.934	0.528	0.728	0.936	0.361	0.868	0.414
	NFH2	0.954	0.358	0.869	0.929	0.358	0.904	0.633
	NOLI	0.940	0.327	0.892	0.936	0.350	0.874	0.719
	POWE	0.942	0.467	0.862	0.935	0.409	0.924	0.804
	TEL2	0.889	0.569	0.695	0.881	0.526	0.834	0.578
	WHIT	0.917	0.162	0.856	0.928	0.333	0.878	0.824
H_O	CLI1	0.880	0.240	0.800	0.880	0.320	0.920	0.800
	CLI2	0.923	0.538	0.692	1.000	0.385	1.000	0.615
	EMOR	0.792	0.708	0.750	0.917	0.125	0.917	0.575
	LITR	0.920	0.333	0.840	0.920	**0.200**	0.960	0.600
	LPIG	0.958	0.583	0.708	0.917	0.458	0.958	0.417
	NFH2	1.000	0.350	0.900	0.850	0.450	0.900	0.580
	NOLI	0.960	0.400	0.860	**0.667**	0.200	0.880	0.700
	POWE	0.917	0.500	0.833	1.000	0.333	0.917	0.667
	TEL2	0.769	0.692	0.769	0.846	0.385	0.769	0.308
	WHIT	0.913	0.174	0.783	1.000	0.227	0.913	0.783
Alleles	CLI1	19	3	6	16	2	14	10
	CLI2	16	3	6	14	3	16	8
	EMOR	16	3	8	22	2	10	5
	LITR	20	3	8	19	3	12	6
	LPIG	17	3	6	16	2	12	4
	NFH2	21	2	8	14	2	13	4
	NOLI	17	2	9	16	2	8	2
	POWE	15	3	8	14	3	13	6
	TEL2	12	3	7	16	3	7	4
	WHIT	21	2	10	16	2	11	10

Table H2, continued

M	CLI1	0.613	1.000	0.667	0.800	1.000	1.000	0.769
	CLI2	0.552	1.000	0.545	0.609	1.000	1.000	0.615
	EMOR	0.615	1.000	0.800	0.786	1.000	1.000	0.625
	LITR	0.769	0.750	0.471	0.826	0.750	0.923	0.600
	LPIG	0.773	1.000	0.750	0.727	1.000	0.857	1.000
	NFH2	0.778	1.000	0.800	0.875	1.000	0.765	0.444
	NOLI	0.850	1.000	0.529	0.727	1.000	1.000	1.000
	POWE	0.577	1.000	0.800	0.583	1.000	0.684	0.857
	TEL2	0.571	1.000	0.778	0.842	0.750	0.875	1.000
	WHIT	0.700	0.750	0.500	0.800	0.750	0.646	0.769

APPENDIX I: Land-use characteristics of darter sampling sites in the upper Tennessee River basin.

Table I1. Land-use characteristics summarized for U.S. Geological Survey 12-digit hydrologic units (HU12s) that contained sites sampled for *Etheostoma rufilineatum* and *E. blennioides* in the upper Tennessee River basin. Site codes correspond to those presented in Table 4.1. Variables are summarized in detail in the text.

Site code	HU12	Cultivated crop area (%)	Pasture area (%)	Developed area (%)	Forested area (%)	Impervious area (%)	Protected area (%)	Road density (km ha^{-1})	NPDES permits (#)	Human population size (thousands)	Human population growth rate
BIGC	060101040102	0.8	18.5	1.9	60.2	0.8	0.0	15.2	0	83.8	0.31
BMOC	060101010402	0.7	31.0	2.2	58.0	1.2	0.0	20.2	2	122.2	0.14
BSYC	060102050901	0.0	10.7	1.2	75.9	0.6	0.0	8.9	0	36.6	0.37
BULL	060102070102	0.1	18.2	5.4	57.4	3.1	0.0	21.6	6	471.2	0.48
CLI1	060102050804	0.0	10.0	3.6	73.3	0.6	0.0	16.1	0	107.3	0.19
CLI2	060102050401	0.6	30.1	0.8	59.9	1.5	0.1	15.7	3	57.3	-0.09
COPP	060102050702	0.8	49.5	0.3	40.0	0.5	0.0	18.9	2	23.4	-0.09
EMOR	060102080402	0.1	6.2	1.4	77.8	0.9	0.0	12.7	0	19.8	0.38
FLAT	060101040306	0.2	28.1	3.2	48.1	1.7	0.9	18.5	3	420.5	0.55
LAUR	060101010105	0.2	9.3	0.0	87.2	0.1	40.5	6.8	0	114.9	0.06
LICK	060101080806	7.5	54.8	2.9	27.5	1.5	1.1	19.6	4	174.6	0.65
LITT	060102010105	1.6	26.3	0.4	66.0	0.4	15.4	18.5	0	177.0	1.16
LPIG	060101070305	0.1	13.6	0.2	81.1	0.3	57.9	18.3	0	71.2	1.93
MFHR	060101020306	2.0	60.4	2.3	26.5	1.4	0.0	19.1	1	101.6	0.18
NFH1	060101010301	0.2	20.3	0.3	75.9	0.2	0.5	11.2	1	98.8	0.21
NFH2	060101010201	0.9	23.7	0.4	70.3	0.3	9.4	11.4	0	101.6	0.18
NOLI	060101080906	2.1	26.8	0.4	61.9	0.5	0.0	18.9	0	96.5	0.47
POWE	060102060304	0.0	24.1	1.5	46.4	0.8	0.2	15.1	0	30.4	-0.10
SFHR	060101020204	0.7	36.8	0.1	56.9	0.3	23.1	16.7	1	239.0	0.33
TEL1	060102040307	0.7	10.3	0.0	82.4	0.1	22.0	15.9	0	39.0	0.67
TEL2	060102040307	0.7	10.3	0.0	82.4	0.1	22.0	15.9	0	39.0	0.67
WACK	060102060302	0.0	8.5	0.9	66.3	0.6	0.1	11.8	1	53.8	-0.09
WHIT	060102010403	0.7	15.6	6.4	56.8	3.6	6.4	23.8	1	80.3	0.46

APPENDIX J: Genetic differentiation between darter populations sampled in the upper Tennessee River basin

Table J1. Estimates of genetic differentiation (F_{ST}) between sites sampled for *Etheostoma rufilineatum* (below diagonal) and *E. blennioides* (above diagonal). Dashes indicate sites not sampled for a given species. Site codes correspond to those presented in Table 4.1.

Site code	BIGC	BMOC	BSYC	BULL	CLI1	CLI2	COPP	EMOR	FLAT	LAUR	LICK	LITT	LPIG
BIGC		-	-	-	-	-	-	-	-	-	-	-	-
BMOC	0.072		-	-	-	-	-	-	-	-	-	-	-
BSYC	0.054	0.092		-	-	-	-	-	-	-	-	-	-
BULL	0.066	0.078	0.047		-	-	-	-	-	-	-	-	-
CLI1	0.039	0.052	0.024	0.025		0.004	-	0.069	-	-	-	0.033	0.052
CLI2	0.063	0.067	0.034	0.026	0.004		-	0.077	-	-	-	0.051	0.075
COPP	0.033	0.069	0.038	0.030	0.010	0.026		-	-	-	-	-	-
EMOR	0.076	0.106	0.050	0.056	0.031	0.039	0.061		-	-	-	0.043	0.023
FLAT	0.053	0.056	0.078	0.067	0.041	0.076	0.020	0.102		-	-	-	-
LAUR	0.091	0.116	0.070	0.046	0.045	0.058	0.062	0.058	0.083		-	-	-
LICK	0.050	0.101	0.056	0.061	0.042	0.061	0.054	0.075	0.077	0.077		-	-
LITT	0.052	0.131	0.047	0.073	0.055	0.062	0.036	0.066	0.074	0.078	0.088		0.045
LPIG	0.035	0.089	0.053	0.035	0.036	0.051	0.025	0.072	0.050	0.080	0.034	0.045	
MFHR	0.027	0.056	0.020	0.033	0.017	0.030	0.037	0.046	0.060	0.063	0.026	0.058	0.028
NFH1	0.061	0.092	0.067	0.039	0.023	0.048	0.040	0.057	0.063	0.004	0.061	0.066	0.058
NFH2	0.073	0.106	0.075	0.041	0.041	0.059	0.061	0.070	0.077	-0.006	0.064	0.086	0.069
NOLI	-	-	-	-	-	-	-	-	-	-	-	-	-
POWE	0.071	0.097	0.053	0.032	0.021	0.026	0.040	0.024	0.076	0.032	0.044	0.060	0.047
SFHR	0.035	0.048	0.036	0.040	0.019	0.032	0.043	0.064	0.059	0.071	0.042	0.079	0.036
TEL1	0.055	0.078	0.074	0.057	0.061	0.088	0.038	0.124	0.055	0.108	0.065	0.088	0.042
TEL2	0.068	0.077	0.089	0.068	0.060	0.089	0.043	0.124	0.054	0.110	0.076	0.104	0.057
WACK	0.045	0.071	0.027	0.035	0.018	0.040	0.022	0.061	0.042	0.065	0.047	0.052	0.032
WHIT	0.056	0.082	0.052	0.056	0.031	0.033	0.045	0.043	0.076	0.057	0.071	0.045	0.055

Table J1, continued.

Site code	MFHR	NFH1	NFH2	NOLI	POWE	SFHR	TEL1	TEL2	WACK	WHIT
BIGC	-	-	-	-	-	-	-	-	-	-
BMOC	-	-	-	-	-	-	-	-	-	-
BSYC	-	-	-	-	-	-	-	-	-	-
BULL	-	-	-	-	-	-	-	-	-	-
CLI1	-	-	0.032	0.036	0.002	-	-	0.054	-	0.011
CLI2	-	-	0.061	0.069	-0.004	-	-	0.077	-	0.031
COPP	-	-	-	-	-	-	-	-	-	-
EMOR	-	-	0.037	0.067	0.058	-	-	0.065	-	0.065
FLAT	-	-	-	-	-	-	-	-	-	-
LAUR	-	-	-	-	-	-	-	-	-	-
LICK	-	-	-	-	-	-	-	-	-	-
LITT	-	-	0.024	0.031	0.030	-	-	0.067	-	0.032
LPIG	-	-	0.019	0.042	0.050	-	-	0.063	-	0.060
MFHR		-	-	-	-	-	-	-	-	-
NFH1	0.038		-	-	-	-	-	-	-	-
NFH2	0.052	-0.006		0.013	0.027	-	-	0.038	-	0.034
NOLI	-	-	-		0.036	-	-	0.069	-	0.036
POWE	0.039	0.025	0.036	-		-	-	0.040	-	0.012
SFHR	0.012	0.048	0.062	-	0.041		-	-	-	-
TEL1	0.066	0.079	0.093	-	0.090	0.068		-	-	-
TEL2	0.072	0.078	0.097	-	0.091	0.062	-0.001		-	0.071
WACK	0.036	0.053	0.055	-	0.046	0.040	0.043	0.058		-
WHIT	0.044	0.051	0.061	-	0.045	0.039	0.085	0.087	0.054	