MIAMI UNIVERSITY
The Graduate School

Certificate for Approving the Dissertation

We hereby approve the Dissertation

of

Emy M. Monroe

Candidate for the Degree:

Doctor of Philosophy

_____

Director
Dr. David J. Berg

_____

Reader
Dr. Thomas O. Crist

_____

Reader
Dr. Brian Keane

_____

Dr. Teresa J. Newton

_____

Graduate School Representative
Dr. Linda E. Watson

Abstract

Species distributions and population genetic structure are determined by landscape history and biological characteristics of an organism. Pleistocene glaciations affected species distributions and freshwater drainage patterns. Extensive work on the biogeography of fish has determined the effects of glaciations on genetic population structure and described recolonization routes from glacial refugia. Freshwater mussels are intimately tied to fish because larvae are largely obligate parasites on fish. Dispersal via hosts has likely led to similar patterns in genetic structure and distributions of mussels and fish. I quantified genetic diversity using mitochondrial DNA and allozymes to determine genetic population structure and infer evolutionary history for *Cumberlandia monodonta* — an endangered species (Chapter 2). There was a lack of genetic structure and evidence of high gene flow. The phylogeny indicated isolation in two glacial refugia followed by post-Pleistocene dispersal of two lineages range-wide. The distribution of *Lampsilis cardium,* a non-endangered species, spans glaciated and non-glaciated areas. Genetic diversity was quantified in northern and southern populations to determine if this large-river species has similar patterns found in other species of fish and mussels (Chapter 3). Genetic data from populations in the upper midwest and Canadian drainages were used to test hypotheses of dispersal from the upper Mississippi River (UMR) into the Red River of the north (RRN). Overall genetic diversity was high, but populations in the north and south had similar levels. There was evidence for two glacial refugia in the Great Lakes basin, but not the upper midwest. Genetic data did not support the hypothesis of post-Pleistocene dispersal from the UMR into the RRN. On a smaller spatial scale, genetic variation and structure of *L. cardium* from the UMR was also quantified with microsatellite markers (Chapter 4). The UMR is fragmented with locks and dams creating distinct river reaches, therefore, population genetic variation and structure were measured within a spatial hierarchy. Historic and contemporary gene flow were compared to determine if the dams are affecting genetic diversity. Genetic diversity was high, and populations in the UMR may be panmictic. There is evidence that fragmentation may be reducing gene flow for this species.

POPULATION GENETICS AND PHYLOGEOGRAPHY OF TWO LARGE-RIVER FRESHWATER MUSSEL SPECIES AT LARGE AND SMALL SPATIAL SCALES

A DISSERTATION

Submitted to the Faculty of

Miami University in partial

fulfillment of the requirements

for the degree of

Doctor of Philosophy

Department of Zoology

by

Emy M. Monroe

Miami University

Oxford, Ohio

2008

Dissertation Director: Dr. David J. Berg

Table of Contents

List of Tables

Chapter 2

Chapter 3

Chapter 4

List of Figures

Chapter 2

Chapter 3

Chapter 4

Dedication

To my husband, Tony: thank you for your unconditional love and support. You were willing to move to Ohio with me to undertake this adventure, and all the while you worked hard to take care of me, the dogs, and provided for our first house! I could not have met this goal without you as my best friend and soul mate.

To my parents, Bruce and Linda Plakke: from day one, you were my first teachers and instilled within me a love of nature. Thank you for your love and guidance throughout my life.

To Teresa Newton: if you had not hired me in 1996 and introduced me to mussels, I would not be where I am today! Thanks for being my first professional supervisor, a great friend, a wonderful mentor, and an inspiration to get my doctorate. I look forward to working as colleagues in the future.

Acknowledgment

I could not have completed this work without the help of many people. I have to thank first and foremost Dave Berg, my major advisor, who was a wonderful and helpful mentor through this process. Dave always encouraged me to participate fully in all aspects of graduate school, which really prepared me for the future. I also thank my advisory committee: Linda Watson, Brian Keane, Tom Crist, and Teresa Newton. I appreciate their guidance and help with this project. Thanks to Linda for hosting several wonderful seminars on systematics and evolution. Thanks to Brian for loaning me lab equipment and for help with laboratory analyses. Tom was one of the best teachers I've had, and really advanced my understanding of ecology. Teresa for supplying the first year of tissue samples and let me stay with her family during field sampling. So many folks helped me in the field: Steve Ahlstead, Dennis Anderson, Shelly Bartsch, Tony Brady, Alan Christian, Jacob Culp, Mike Davis, Cathlene and Curt Elderkin, Scott Faiman, John Harris, Randy Hines, Mark Hove, Jess Jones, Bob Kennedy, Todd Levine, Monte McGregor, Jenny Sauer, Bernard Sietman, Daelyn Woolnough, and Dave Zanatta. The USGS Upper Midwest Environmental Sciences Center provided vehicles and boats for field work on the Upper Miss. Joseph Carney kindly donated all of the samples from Canadian rivers and I am so very grateful! I appreciate museum donations of tissue clips from Art Bogan of the University of North Carolina Museum, Kevin Cummings of the Illinois Natural History Survey, Dan Graf of the Academy of Natural Sciences in Philadelphia, Tim Pierce of the Carnegie Museum of Natural History, and Tom Watters and Sarah Hazzard of the Ohio State University Museum of Biological Diversity. I also appreciate help in the laboratory from several undergraduate students over these last five years: Nancy Benight, Nikki Ferreira, Allison Fralick, Jessica Hoisington, Obie Iloka, Jessica Janus, Jorge Lassus, and Marnin Wolfe. Not only were they great help, but good friends too. My lab group was very helpful in discussing ideas and reading versions of these chapters: Curt Elderkin, Todd Levine, Makiri Sei, and Rick Seidel. Chris Wood and the technicians in the Center for Bioinformatics and Functional Genomics at Miami University were so helpful and made collecting data smooth and easy! Sherry House took care of computers in the lab and my office with a smile. I was really lucky to share an office for four years with Jim Stoeckel, he was a great friend and made hours in the office fun. I was lucky to share the office with Jen Purrenhage and Lesley Knoll during my last year; I am glad we could happily coexist in such a small space! Finally, thanks to the entire Department of Zoology faculty, students, and staff, a genuinely great group of people!

Funding for chapter 2 was provided by the US Fish and Wildlife Service, and all other work was supported by the Zoology Department, a Doctoral-Undergraduate Opportunities for Scholarship grant, and the Graduate School.

Chapter 1

Introduction

Riverine landscapes are dynamic ecosystems that are biologically and spatially complex; as a result they generate and maintain high levels of biodiversity (Robinson *et al.* 2002). Human modification of landscapes and river channels has degraded freshwater ecosystems on a global scale (Nilsson *et al.* 2005). Rivers and streams have been altered by the construction of dams, resulting in loss and fragmentation of lotic habitat (Watters 2000). All of the large rivers (> 1000 km long) in the contiguous US have been altered for navigation or hydropower (Benke 1990), which has negatively affected the biodiversity supported by freshwater systems. Projected mean future extinction rates for North American freshwater fauna were calculated to be about five times greater than the rate for terrestrial fauna and three times the rate for coastal marine mammals (Ricciardi & Rasmussen 1999). It is generally accepted that biota on this planet are in the midst of a sixth massive extinction (Leaky & Lewin 1995), and freshwater systems contain some of the most highly endangered faunas (Higgins *et al.* 2005). Molecular genetic studies can provide information that is useful in guiding conservation of endangered freshwater fauna, while advancing our understanding of evolution and how various processes extrinsic and intrinsic to organisms affect genetic diversity (Lowe *et al.* 2003).

Evaluating large-scale landscape-level effects in a phylogeographic context advances our understanding of the evolutionary history of species and can provide guidance for conservation measures. For example, when the evolutionary history within a species indicates long-term separation of populations, then conservation efforts should maintain these separate evolutionary trajectories, commonly called evolutionary significant units (ESU's; Moritz 1994). Conservation of ESU's in general considers long-term conservation goals, mainly because it relates to conserving a species' evolutionary potential (Moritz 1994). In contrast, short-term conservation goals address threatened populations that have suffered reduced population size or loss of genetic variation (Moritz 1994), which is often directly related to human activity. Human induced changes to the landscape, such as the impoundment of rivers in North America, are recent

compared to Pleistocene climate change. Short-term conservation in this context requires information on genetic diversity and current levels of gene flow among extant populations.

The most imperiled faunal group in North America is freshwater mussels (Richter *et al.* 1997), which is also the most diverse assemblage of freshwater mussels in the world (Bogan 1993). Therefore, many federal, state, and local agencies are examining approaches to conserve this declining fauna (Newton *et al.* 2001). Conservation tools include artificial propagation and relocation (Cope & Waller 1995; U.S. Fish & Wildlife 1997, 2003). Propagation is being employed for critically endangered species and relocation is often utilized as a way to address population declines due to zebra mussel infestation or to mitigate direct impacts from the construction of bridges or navigational structures or other habitat fouling or loss. However, effective conservation of these organisms requires not only propagating and saving individuals, but also maintaining natural levels of genetic diversity.

Genetic diversity maintains a species' evolutionary potential and capacity for adaptation to environmental changes over time. For small populations, genetic diversity is needed because it contributes to fitness-related traits such as survivorship and fecundity (Mulvey *et al.* 1998). Typical management actions such as propagation and augmentation may reduce overall genetic diversity if the wild genotypes are swamped by less diverse genotypes of propagules (Frankham 2003). Similarly, relocating mussels between genetically distinct populations can also reduce overall diversity by artificially mixing historically isolated populations, leading to introgression and/or hybridization (Frankham 2003). As a result, it is critical that genetic structure of endangered species be understood before management actions are begun. The dearth of information on genetic diversity in freshwater mussels is an obstacle to implementing effective conservation strategies for this faunal group (National Native Mussel Conservation Committee 1998).

The life cycle of mussels is complex, which makes studying these organisms challenging, and also likely contributes to their imperiled status. These animals are very long-lived, with some species living around a decade and others surviving up to a century (Anthony *et al.* 2001). Similar to other long-lived species, reproduction is

delayed as long as 13 years (Young & Williams 1984). Larval mussels (glochidia) are brooded in the gills of female mussels, and in most cases, must attach to a host fish for transformation into the juvenile life stage. This intimate relationship with fish hosts has implications for mussel dispersal. Adult mussels are relatively sessile, but glochidia can be dispersed upstream and downstream of their natal mussel bed by their hosts. Some mussel species are able to transform on several species of fish, while others are restricted to only a one or a few species of fish (Mussel Host Database, The Ohio State University Museum of Biological Diversity, http://www.128.146.250.63/bivalves). Depending on which fish host a mussel species uses, some mussels will be dispersed greater distances than others, and therefore, may have different levels of gene flow among populations.

Freshwater mussels are also ideal organisms for ecological genetic studies for several reasons. They live in linear river systems connected within a nested hierarchy. Therefore, genetic diversity can be quantified along sliding spatial scales from the individual or population level to increasingly larger distances (Excoffier *et al.* 1992). Variation over regional spatial scales can assess the relative importance of gene flow and genetic drift (Hutchison & Templeton 1999). Understanding the influences of these two genetic processes is useful for interpreting historical evolutionary events and understanding current gene flow over the landscape.

Extensive work on the zoogeography of fish in North America has created several phylogeographic hypotheses to explain both species distribution and population genetic structure (Mayden 1988; Wiley & Mayden 1985). Comparative phylogeographic studies for several species have led to generalizations about the affects of Pleistocene glaciations on terrestrial and aquatic fauna (Avise 2000; Bernatchez & Wilson 1998). Studies on fish have commonly shown that populations north of glacial maxima have less genetic diversity as result of a recent founder effects from post-Pleistocene migration and colonization northwards compared to southern populations with their longer histories (Bernatchez & Wilson 1998). Another common finding is evidence for multiple glacial refugia for terrestrial (Spellman & Klicka 2006) and aquatic species (Bernatchez 2001). Because of their relationship with fish, mussels can be used to test hypotheses generated from the body of work on fish phylogeography.

In general, phylogeographic studies of mussels have shown similar patterns to those found in fish; mussel populations north of glacial maxima have lower genetic diversity than those in the south (Elderkin *et al.* 2007, 2008; Zanatta & Murphy 2008). Similar to fish, mussel populations in the north were colonized from multiple glacial refugia (Zanatta & Murphy 2008; Elderkin *et al.* 2008). New analytical methods such as statistical phylogeography (Knowles 2004) can explicitly test hypotheses about the number and location of glacial refugia. These methods have been applied to terrestrial species (Spellman & Klicka 2006), but as far as I am aware, not yet to aquatic species.

My research will add to the understanding of genetic diversity in freshwater mussel species at large spatial scales over evolutionary history. In chapters 2 and 3, two species are studied within a phylogeographic context to understand how Pleistocene glaciations and subsequent gene flow has affected population genetic structure of mussels in a large river system. Mitochondrial DNA (mtDNA) was used to measure genetic diversity and population differentiation at medium-to-large spatial scales (Lowe *et al.* 2004). Chapter 2 is focused on a critically imperiled species, *Cumberlandia monodonta*, while chapter 3 looks at a relatively healthy and widespread species, *Lampsilis cardium*. These two chapters address the evolutionary histories of mussel species, but effects on dispersal and gene flow among extant populations from recent anthropogenic affects are not discernable with mtDNA markers (Moritz 1994).

Current levels of gene flow and population dynamics in aquatic invertebrate species have been greatly affected by impoundments due to several reasons (Strayer 2006). One of the most obvious effects on freshwater mussels is habitat loss which creates fragmentation among populations in both small and large streams (Watters 2000). Dams are effective barriers against fish movement (Watters 2000), which in turn prevents mussel dispersal and gene flow. Gene flow is an important process for maintaining genetic diversity (Templeton *et al.* 1990) and connectivity among populations. Application of genetic studies may be able to determine if dams are reducing genetic diversity in mussel populations. In fish, dams have been shown to reduce genetic diversity in some species, but not others (Reid *et al*. 2008). To date, most population genetic studies on mussels have been on already-threatened populations (Bowen & Richardson 2000; Kelly & Rhymer 2005; Zanatta & Murphy

2008), and no one has looked at fine-scale genetic diversity in a highly fragmented system in North America (but see Geist & Kuehn 2005).

My work in chapter 4 is a small-scale geographic study designed to measure fine-scale genetic structure with highly variable microsatellite markers. The highly fragmented upper Mississippi River is dominated by a system of locks and dams that create distinct river reaches. If habitat fragmentation has created barriers to gene flow among populations of mussels, then genetic diversity should be partitioned within and among river reaches and there should be differences in historic versus contemporary gene flow. Gathering this information with a relatively abundant species of mussel will provide insight into the natural genetic structure and gene flow of mussels in large rivers, something that is not attainable with endangered species that are not populous enough to provide robust sample sizes.

References

Anthony JL, Kesler DH, Downing WL, Downing JA (2001) Length-specific growth rates in freshwater mussels (Bivalvia: Unionidae): extreme longevity or generalized growth cessation? *Freshwater Biology,* **46**, 1349-1359.

Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, Massachusetts.

Benke AC (1990) A perspective on America's vanishing streams. *Journal of the North American Benthological Society,* **9**, 77-88.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Neararctic and Palearctic fishes. *Molecular Ecology*, **7**, 431-452.

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

Bogan AE (1993) Freshwater bivalve extinctions (Mollusca, Unionoida) - a search for causes. *American Zoologist,* **33**, 599-609.

Bowen BS, Richardson WR (2000) Genetic characterization of *Lampsilis higginsii*. Final report to the US Fish and Wildlife Service, Bloomington, MN.

Cope WG, Waller DL (1995) Evaluation of freshwater mussel relocation as a conservation and management strategy. *Regulated Rivers: Research and Management*, **11**, 147-155.

Elderkin CE, Christian AD, Vaughn CC, Metcalfe-Smith JL, Berg DJ (2007) Population genetics of the freshwater mussel, *Amblema plicata* (Say 1817)(Bivalvia: Unionidae): evidence of high dispersal and post-glacial colonization. *Conservation Genetics*, **8**, 355-372.

Elderkin CE, Christian AD, Metcalfe-Smith JL, Berg DJ (2008) Population genetics and phylogeography of freshwater mussels in North America, *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia: Unionidae). *Molecular Ecology*, **17**, 2149-2163.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491

Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies,* **326**, S22-S29.

Geist J, Kuehn R (2005) Genetic diversity and differentiation of central European freshwater pearl mussel (*Margaritifera margaritifiera* L.) populations: implications for conservation and management. *Molecular Ecology,* **14**, 425-439.

Higgins JV, Bryer MT, Khoury ML, Fitzhugh TW (2005) A freshwater classification approach for biodiversity conservation planning. *Conservation Biology,* **19**, 432-445.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898-1914.

Kelly MW, Rhymer JM (2005) Population genetic structure of a rare unionid (*Lampsilis cariosa*) in a recently glaciated landscape. *Conservation Genetics,* **6**, 789-802.

Knowles LL (2004) The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*, **17**, 1-10.

Leaky R and Lewin R (1995) *The sixth extinction: biodiversity and its survival.* Doubleday, New York.

Lowe A, Harris S, Ashton P (2004) *Ecological Genetics*. Blackwell Science Ltd., Malden, Massachusetts.

Mayden RL (1988) Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Systematic Zoology*, **37**, 329-355.

Moritz C (1994) Applications of mitochondrial-DNA analysis in conservation - a critical review. *Molecular Ecology*, **3**, 401-411.

Mulvey M, Liu H, Kandl KL (1998) Application of molecular genetic markers to conservation of freshwater bivalves. J*ournal of Shellfish Research,* **17**, 1395-1405.

Newton TJ, Monroe EM, Gutreuter SJ, Kenyon R, Thiel PA, Welke KI (2001) Evaluation of relocation of unionid mussels into artificial ponds. *Journal of the North American Benthological Society,* **20**, 468-485.

Nilsson C, Reidy CA, Dynesius M, Revenga C (2005) Fragmentation and flow regulation of the world's large river systems. *Science,* **308**, 405-408.

Reid SM, Wilson CC, Mandrak NE, Carl LM (2008) Population structure and genetic diversity of black redhorse (*Moxostoma duquesnei*) in a highly fragmented watershed. *Conservation Genetics,* **9**, 531-546.

Ricciardi A, Rasmussen JB (1999) Extinction rates of North American freshwater fauna. *Conservation* Biology, **13**, 1220-1222.

Richter BD, Braun DP, Mendelson MA, Master LL (1997) Threats to imperiled freshwater fauna. *Conservation Biology,* **11**, 1087-1093.

Robinson CT, Tockner K, Ward JV (2002) The fauna of dynamic riverine landscapes. *Freshwater Biology,* **47**, 661-677.

Spellman GM, Klicka J (2006) Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pynmaea*). *Proceedings of the Royal Society B*, **273**, 3057-3063.

Strayer DL (2006) Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society,* **25**, 271-287.

Templeton AR, Shaw K, Routman E, Davis SK (1990) The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden,* **77**, 13-27.

U.S. Fish and Wildlife Service (1997) Winged mapleleaf mussel (*Quadrula fragosa*) recovery plan. Ft. Snelling, Minnesota.

U.S. Fish and Wildlife Service (2003) Higgins' eye pearlymussel (*Lampsilis higginsii*) draft recovery plan: first revision. Ft. Snelling, Minnesota.

Watters GT (2000) Freshwater mussels and water quality: a review of the effects of hydrologic and instream habitat alterations. *Proceedings of the First Freshwater Mollusk Conservation Society Symposium,* 261-274.

Wiley EO, Mayden RL (1985) Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. *Annals of the Missouri Botanical Garden,* **72,** 596-635.

Young M, Williams J (1984) The reproductive biology of the freshwater pearl mussel *Margaritifera margaritifera* (Linn.) in Scotland I. Field studies. *Archiv für Hydrobiologie,* **99**, 405-422.

Zanatta DT, Murphy RW (2008) The phylogeographical and management implications of genetic population structure in the imperilled snuffbox mussel, *Epioblasma triquietra* (Bivalvia: Unionidae). *Biological Journal of the Linnaean Society*, **93**, 371-384.

Chapter 2

Phylogeography reveals Pleistocene isolation followed by high gene flow in a wide-ranging imperiled mussel: the spectaclecase, *Cumberlandia monodonta*

Abstract

Freshwater organisms of North America have had their population genetic structure shaped by past events, such as Pleistocene glaciations. Life history traits that promote dispersal and gene flow have since continued to shape population genetic structure. Freshwater mussel larvae are obligate parasites of fish hosts, and thus, fish movement is the major mode of dispersal for mussels. *Cumberlandia monodonta*, an imperiled mussel with a widespread distribution was examined to determine genetic diversity and population genetic structure range-wide. Two different classes of molecular markers, allozymes and cytochrome oxidase I (COI) partial gene sequences were used to measure genetic diversity. COI data were analyzed with two phylogeographic methods: nested clade analysis and statistical phylogeography. Explicit hypotheses were tested to determine if this species spread northward from two glacial refugia: 1) areas south of the glaciers and, 2) the Driftless Area of WI, IA, MN, and IL, or only one refuge. There was not any geographic structure detected with allozyme or COI data. A phylogeny of COI sequences indicated past isolation which created divergence of two lineages estimated to occur during the Pleistocene. However, isolation did not occur within the driftless area. Up until recent times, gene flow within this species has maintained high levels of genetic diversity. Extirpation of this species from its entire central range likely has isolated remaining populations due to the distances among them.

Introduction

Species distributions and population genetic structure are determined by both landscape history and biological characteristics of an organism (Vucetich & Waite 2003; Wares & Turner 2003; Whiteley *et al.* 2004). Pleistocene glaciations affected extant species distributions more so than any other past event (Roy *et al.* 1996). Organisms survived in suboptimal habitats in near-glacier zones, migrated further south, or were extirpated from glaciated areas. Even southern areas were not free from affects due to climatic and physical changes. The glaciers created a drier, colder climate and more land was exposed due to lowered sea levels (Smith 2007). Many researchers have hypothesized that organisms often survived the Pleistocene in multiple isolated refugia, where small population sizes and different environmental conditions created intraspecific variation (reviewed in Hewitt 1996). These processes left their mark in the genes of species, and can be used to infer past evolutionary and demographic events within species (Avise 2000).

Freshwater systems, and the organisms living in them, were altered by the Pleistocene when rivers were often rerouted (Berendzen *et al.* 2003) and large areas of land were inundated, allowing aquatic organisms to disperse widely and facilitating mixtures of conspecifics from different refugia (Bernatchez & Wilson 1998; Gagnon & Angers 2006). The effects of these events on distribution and population structure of fish have been studied extensively and the signature of Pleistocene glaciations on the evolutionary history of several fish species is apparent from comparative studies. When phylogenies of several fish species with ranges in formerly glaciated and non-glaciated regions are compared, populations in glaciated areas often have lower levels of genetic diversity compared to populations from non-glaciated areas and the former show signs of range expansion (Berendzen *et al.* 2003; Bernatchez & Wilson 1998). Biogeographic analyses of fish (Mandrak & Crossman 1982) and mussels (Graf 1997) have resulted in several different routes of dispersal from multiple refugia into northern systems proposed for both taxa. While freshwater fish phylogeography has been extensively studied, there is not nearly as much work on other freshwater organisms.

Freshwater mussels reach maximum biodiversity in North America (Lydeard *et al.* 2004) and have unique life histories that are intimately tied to fish because larvae

(glochidia) are obligate parasites of fish. Dispersal via fish hosts has likely led to similar patterns in population structure and distributions of mussels and fish. Despite this close relationship, most phylogeographic studies of mussels have been limited to endangered species with either very small population numbers or limited geographic ranges (Buhay *et al.* 2002; King *et al.* 1999; Roe *et al.* 2001). Recently however, Elderkin *et al.* (2007, 2008) examined the phylogeography of several widespread and relatively healthy mussel species that live in the Mississippi and Great Lakes drainages. It appears that mussels colonized northern systems, such as the Great Lakes, upper Mississippi and other Canadian drainages, from multiple refugia (Elderkin *et al.* 2008) as did fish (Bernatchez & Wilson 1998). These northern populations have been shown to have less genetic diversity than those in southern areas that were never glaciated (Bernatchez & Wilson 1998). However, mussel species that co-occur in the same aggregations can have different genetic structures at various spatial scales (Elderkin *et al.* 2008). These differences are likely due to variations in life-history traits, such as use of different fish hosts and demographic differences among species (Elderkin *et al.* 2008). This variation among mussel species becomes important when considering conservation efforts for threatened or endangered species, a status that applies to 70% of all mussel taxa in North America (Lydeard *et al.* 2004). Even though threatened species may live side-by-side with healthy species, if their genetic structures differ, then each species would require different conservation action. Here, I survey the genetic population structure of *Cumberlandia monodonta* (Say), a historically widespread species in the Mississippi River system that occupies both formerly glaciated and non-glaciated areas.

Cumberlandia monodonta, the spectaclecase, is the sole species of the genus within the family Margaritiferidae. Although this family is circumpolar (ORVMTS 2002), *C. monodonta* distribution is limited to the eastern United States. Until the last half of the 20th century, it was found in the mainstems of the Mississippi, Ohio, and Tennessee rivers as well as 40 tributaries; extant populations are limited to a few that are located at the periphery of its former range. It has been sporadically reported in 19 streams over the last 20 years, and decline of this species is likely due to its preference for riffle areas of streams (ORVMTS 2002) and the loss of these habitats due to inundation and

fragmentation from impoundment of rivers (Nilsson *et al*. 2005). In addition to habitat loss, the fragmentation of rivers has had serious ramifications for fish movement and, therefore, mussel dispersal and gene flow. *Cumberlandia monodonta* individuals live up to 56 years and reach sexual maturity around 5-6 years, with maximum reproductive effort starting around 10 years (Baird 2000). This species is dioecious and females release anywhere from 3-9 million glochidia in April or May. Unfortunately, the host fish is still unknown (Baird 2000; Hove *et al*. 1998).

Distribution of *C. monodonta* prior to its decline in the last century was very similar to several fish species that have been studied using molecular genetics and phylogeographic analyses (Bernatchez & Wilson 1998; Near *et al.* 2001; Strange & Burr 1997). Comparisons among fish species, especially those found in previously glaciated areas, have revealed similar changes in overall genetic diversity (Bernatchez & Wilson 1998), but different colonization histories. Fish inhabiting Mississippian refugia were able to use proglacial lakes for dispersal into northern habitats, facilitating many different patterns of dispersal as lakes changed with receding and advancing ice sheets (Gagnon & Angers 2006; Mandrak & Crossman 1992). With similar distributions, *C. monodonta* could have also survived glaciations in multiple southern refugia, as well as in more isolated areas where genetic drift could occur. One such location may have been a northern refugium in the driftless area in southwestern Wisconsin (Gibbard & Van Kolfschoten 2004). The driftless area contains several relict species from the Pleistocene, including land snails, mites, and plants (Nekola 1999; Ross 1999). The extant population of the spectaclecase in the St. Croix River is proximal to the driftless area (Figure 1), and could have been colonized by mussels surviving in streams of this region.

The goals of this study were to measure genetic variation in extant populations of *C. monodonta*, and to determine its evolutionary history since the Pleistocene. The objectives to meet the first goal were to use allozyme and cytochrome oxidase I (COI) partial gene sequences to: 1) quantify within-population variation, 2) quantify among-population variation, and 3) determine how variation is partitioned within and among populations. Hypotheses for these objectives were as follows: 1) within-population genetic variation will be high, but 2) genetic variation will be low among populations, and

3) there will be very low genetic population structure because most of the variation will be contained within populations. An objective to meet the second goal was to compare genetic diversity of a population in the north to those in the south. The northern population is in an area that was previously glaciated. Therefore recent founder effects should have reduced genetic diversity compared to southern populations that were stable for longer time periods which allowed neutral variation to increase genetic diversity. A final objective to meet the second goal used two phylogeographic methods was to evaluate evolutionary history with COI sequences: nested clade analysis (NCA; Templeton *et al.* 1987; Templeton 2004) and statistical phylogeographic methods (Knowles 2001; Richards *et al.* 2007; Smith 2007). Nested clade analysis inferred the effect of past events on genetic diversity, and statistical phylogeography tested if within-species diversity arose from isolation in multiple refugia or one refuge during the Pleistocene. The hypothesis for the first objective to meet the second goal is that genetic diversity in a northern population should be lower than that in southern populations. They hypothesis for the second objective to meet the second goal is that there should be evidence of isolation in multiple glacial refugia in the northern population.

Materials and methods

Collection

Tissue was obtained from a total of 178 individuals (Table1) from 4 extant populations (Figure 1). Museum specimens were also sampled, which added a 5th population in the TN River and a singleton from the Cumberland River (Figure 1). Three large populations were sampled in 2002: 48 individuals from the Clinch River in Hancock County, TN; 48 from the St. Croix River in Chisago County, MN; and 40 from the Gasconade River in Osage County, MO (Figure 1). An additional 34 animals were taken from a fourth large population in the Meramec River, Franklin County, MO in 2004. All live specimens were collected by wading or diving; animals were sampled non-destructively by clipping mantle tissue (Berg *et al.* 1995) and then returning them to the river bottom. Tissue was carefully taken from somatic tissue to avoid the issue of doubly uniparental inheritance (Zouros *et al*. 1994). Tissue samples were flash frozen in liquid nitrogen in 2002 for allozyme and mtDNA, analyses and preserved in 95%

ethanol in 2004 for mtDNA sequencing. A voucher specimen from the Meramec population was placed in the Illinois Natural History Survey Collection (number INHS 31198). Museum specimens were clipped by museum personnel and sent to Oxford, OH. Museum specimens from the Tennessee River were actually collected from two locations 135 km apart; however, due to the large spatial scale of this study, these were considered as one population for all analyses. Three extant populations (Gasconade, MO; Clinch, TN/VA; St. Croix, MN/WI) were analyzed with both allozyme electrophoresis and sequencing of the mitochondrial cytochrome oxidase subunit I (COI) gene, but specimens from all other localities were only sequenced.

Allozymes

Allozymes were visualized using starch gel electrophoresis following standard methods (Hebert & Beaton 1993; Richardson *et al.* 1986) with modifications. Nine enzyme systems revealed a total of 11 putative loci. These loci were aspartate aminotransferase, (AAT), esterase (EST; 1 visible and 1 ultraviolet locus), glucose-6-phosphate isomerase (GPI), isocitrate dehydrogenase (IDH), leucine aminopeptidase (LAP), malate dehydrogenase (MDH), phosphoglucomutase (PGM-Fast and PGM-Slow), peptidase (PEP), and superoxide dismutase (SOD). Each tissue sample was homogenized with 28 µl of 2% 2-phenoxyethanol and then centrifuged at 14,000 x g for 3 min. The supernatant solution was soaked onto filter-paper wicks, and loaded into starch gels and run for various lengths of times depending on the buffer system used (JRP for AAT, IDH, MDH; Tris-citrate 6.7 for EST, LAP, PEP, PGM; Lithium hydroxide for GPI, SOD). Alleles were sequentially numbered from the fastest to the slowest migration rate towards the anode.

mtDNA

Total genomic DNA was extracted from recently preserved tissues using ~25 mg tissue and DNeasy animal tissue kits (Qiagen, Valencia, CA). Similar quantities of tissue from museum specimens were extracted with phenol/chloroform (Sambrook *et al.* 1989). Polymerase chain reaction and sequencing was performed as in Elderkin *et al.* (2007), except BigDye v3.1 (Applied Biosystems Inc, Foster City, CA) was used for sequencing reactions. Sequences were easily aligned by eye in BioEdit 7 (Hall 1999)

because there were no insertions or deletions, and base calls were confirmed by checking all polymorphic sites in the raw electropherograms.

Data Analysis

Allozymes

Analyses of electrophoresis results were conducted using BIOSYS-1 (Swofford & Selander 1981) and TFPGA (Miller 1997) to obtain descriptive measures of within-population genetic variation. Exact tests of goodness-of-fit determined whether genotype frequencies met Hardy-Weinberg expectations. A one-way ANOVA was used to test differences in average heterozygosity among populations. Among population variation was described by comparing allele frequencies among populations for polymorphic loci and evaluated with Nei's (1978) unbiased genetic distance (D) and modified Rogers distance (Wright 1978). Differentiation among populations was quantified with exact tests and Weir and Cockerham's (1984) method of calculating Wright's (1965) F-statistics (θ).

mtDNA

Genetic diversity was analyzed for haplotype frequencies and molecular divergence in Arlequin 3.01 population genetic data software (Excoffier *et al.* 2005). Arlequin was also used to calculate mean nucleotide diversity (π) within each population as a measure of variation within populations. The number of haplotypes, nucleotide diversity, and the mean number of basepairs different within populations were tested for correlation with latitude using linear regression analysis. The regression analyses only included the 4 large extant populations, because sample sizes were similarly large, the Tennessee population did not have the same sample size. Genetic differentiation was analyzed within and among all five populations (AMOVA; Excoffier *et al.* 1992), and significance of this differentiation was tested using 1023 nonparametric permutations (Weir & Cockerham 1984). The mean value of F_{ST} differences among pairs of populations was calculated with 1023 permutations to test significance at $p = 0.05$. Matrices of pairwise genetic differences and geographic distances between populations were evaluated for significant correlation with a Mantel test using Arlequin. The relative contributions of gene flow and genetic drift in determining population genetic structure

(Hutchison & Templeton, 1999) were evaluated by a manual Mantel test in Arlequin using matrices of residual pairwise F_{ST} values and geographic distance. Geographic distances between each pair of populations were measured to the nearest km using the path measurement tool in GoogleEarth© (www.googleearth.com) by following each bend in the river. Populations were placed into GoogleEarth with coordinates obtained onsite with various handheld Global Positioning Units.

Phylogenetic relationships were analyzed with Bayesian inference, maximum parsimony (MP), and maximum likelihood (ML) and included *Margaritifera auricularia, M. margaratifera, M. falcata, and M. laevis* sequences from Huff *et al.* (2004), and one unionid *Anodonta* spp. sequence from GenBank (AY493505) as outgroups. MrModeltest (version 2.2, Nylander 2004) was used to determine the most suitable model of DNA evolution for ML and Bayesian analyses with a hierarchical log-likelihood-ratio test. The best evolutionary model for the data set was the GTR+Γ+G model, and this was used for ML and Bayesian analyses. Bayesian inference was conducted in MrBayes v3.0 (Huelsenbeck & Ronquist 2001). Two runs of 4 chains with default heat settings were conducted with a random starting tree for 1.6 million generations, sampling trees every 100 generations and excluding the first 200 trees as burn-in. The number of trees to exclude as burn-in was evaluated by visualizing when the Markov chain Monte Carlo (MCMC) analysis reached stationarity (Hulesenbeck & Ronquist 2001). A 50% majority rule consensus tree was generated from 15,800 trees and the percentage of time each node occurred was reported as the posterior probability of the node (Hulesenbeck & Ronquist 2001). MP was performed in *MEGA* version 4 (Tamura *et al.* 2007) with a heuristic branch swapping search and the initial tree generated by random addition. Bootstrap (1000) repetitions (Felsenstein 1985) were performed to estimate support for the most parsimonious tree. ML inference was carried out using phyML (version 2.4.4, Guindon & Gascuel 2003), with 1000 bootstraps to estimate nodal support.

A 90% confidence parsimony network was created with TCS version 1.13 (Clement *et al.* 2000) and multiple connections between haplotypes were simplified using an approach as outlined in Fetzner and Crandall (2003). Mutations were considered more probable if they arose from more frequent haplotypes; those

haplotypes with more than one pathway were assigned the shortest path chosen from the most frequent haplotype. Haplotype frequencies were summed based on network relationships and placed on the site map. Nested clade analysis was conducted in Geodis (Posada *et al.* 2000, 2006) after hand-nesting haplotypes in the network according to methods described by Templeton (1998), with modified rules from Templeton and Sing (1993). Output of the Geodis software was interpreted with the inference key of Templeton (1998, 2004). A mismatch distribution was also calculated in Arlequin for each lineage, for each population, and for all of the individuals pooled.

Coalescent hypothesis testing

Coalescent hypothesis testing is described in detail elsewhere (Richards *et al.* 2007), but briefly, it requires two inputs: a set of data simulated under alternate historical population hypotheses, and an empirical genetic data set. Each replicate of the simulated data can be characterized using a summary statistic, which generates an expected pattern of genetic variation under each historical scenario. The same summary statistic is also computed for the empirical genetic data and compared with the simulated distributions of that statistic. Each of these components is a multi-step procedure and will be described below.

Alternate hypotheses of dual-refugia and single-refugium were tested for this species. A hypothesis of two refugia was tested (Figure 2a) because it is consistent with the phylogeography of several freshwater fish (Bernatchez & Wilson 1998; Mandrak & Crossman 1988) and mussel species (Graf 1997; Zanatta & Murphy 2008). The dual-refugia hypothesis included one refuge south of the glacial maximum and a second refuge in the Driftless Area. To model this refuge, the St. Croix River population closest to this region was used (Figure 1). The alternate single-refugium hypothesis tested was that these mussels survived only in populations south of the glacial maximum (Figure 2b).

Hypotheses were tested with the coalescent module (Maddison 2007) and simulations were conducted in Mesquite v. 1.05 (Maddison & Maddison 2007), where 1000 coalescent genealogies (with migration) were generated under each historical population scenario. To test each hypothesis, a statistical measure of the gene tree must be compared to a distribution of the same statistic for gene trees simulated onto

population hypotheses (Richards *et al*. 2007). In this case, because high levels of gene flow dispersed haplotypes among all of the populations, Slatkin and Maddison's (1989) S-value was chosen as the test statistic. S is the number of migration events that had to have occurred to distribute haplotypes among locations where they were found. Failure to reject an historical population model occurs when the value of *S* for the empirical gene tree falls within the 95% confidence intervals of the simulated distribution of S-values.

To simulate genetic data onto each historical population model in Mesquite, estimates of time and effective population size (N_e) were required. To determine a reference time frame to test each hypothesis, the time of divergence of lineage 2 from lineage 1 was roughly estimated using the method of Sei and Porter (2007). This absolute time in years was converted to generation times with the minimum of 6 and a maximum of 26 years. Two generation times were used to cover all reasonable estimates of time for a long-lived species such as *C. monodonta*. A mean generation time of 26 years was calculated by constructing a static life table with demographic data (Baird 2000), and a minimum generation time was considered as the age of first reproduction, which is 6 years (Baird 2000).

Simulations also require an estimate of N_e, which was calculated from θ (the number of heritable haplotypes in a population), a value generated in the coalescent program Lamarc (Kuhner 2006). Bayesian analysis was used to estimate θ with one long chain of 50,000 steps sampled every 100 steps following a burn-in of 1500 steps. Lamarc was run four times to ensure convergence with different random seed numbers; curve and trace files were used to ensure sufficiently long chains were run (Kuhner & Smith 2007). Once θ values were obtained, N_e was calculated with $\theta = 2N_e\mu$, where the mutation rate (μ) was 3.16×10^{-7} or 1.37×10^{-6} substitutions per site per generation (COI specific rate from bivalves; Marko 2005), depending on the generation time used (26 or 6 years respectively).

Lamarc (Kuhner 2006) was also used to estimate immigration rates from each population into all other populations. This was done with four separate Bayesian analyses, each consisting of a single long chain of 50,000 steps sampled every 100 steps and all trees discarded before 1000 burn-in steps. The Tennessee population

was eliminated from this analysis because Lamarc documentation recommends only using populations with 10 or more individuals. Again, curve and trace files were used to evaluate whether chains were run long enough to reach stationarity (Kuhner & Smith 2007).

Results

Allozymes

Using a 95% criterion, 3 of 11 loci were polymorphic in all populations, and an additional locus was polymorphic in only one population. Mean sample size per polymorphic locus was 45.5 individuals, with an average of 1.91 alleles per locus; populations had a range of 27-36% polymorphism (Table 2). Of the 7 loci that were not polymorphic, four of them were fixed for a single allele in all populations: both EST loci, MDH, and SOD (data not shown). Mean heterozygosity was not different among populations (df = 1,31, F = 0.52, p = 0.48) and averaged 0.145 (range 0.13-0.16). Exact tests of allele frequencies did not vary among populations for the three polymorphic loci (PGM-F, LAP, and PEP; df = 14, X^2 = 13.1, p = 0.52; Table 2). Nei's unbiased genetic distances among pairs of populations were zero and Rogers's genetic distances were all 0.03. Additionally, Weir and Cockerham's θ was not different from zero, with 95% confidence intervals surrounding zero (-0.0286 to 0.0966).

mtDNA

A 678-bp region of the COI gene was sequenced in 178 individuals from 5 populations (plus a singleton) which revealed 30 unique haplotypes with 46 phylogenetically informative sites. The Clinch River population had the greatest variation with 18 haplotypes, and each of the other four populations contained six to nine haplotypes (Table 1). Unique haplotypes were found in each population; the Clinch population had the most with 11 (Table 1). Nucleotide diversity values were extremely low, ranging from 0.0015 to 0.0073 (Table 1). There were no significant correlations of diversity with latitude. The Cumberland River haplotype from a museum specimen was one of the most common haplotypes in other populations.

Analysis of molecular variance (AMOVA) revealed 95% of genetic variation was within populations and all sequences were only 5.3% (±0.3 SE) divergent. Pairwise

comparisons of each sequence form a bi-modal mismatch distribution (Figure 3a) with distinct modes at 3 and 8 bp; these modes are also present in each population mismatch distribution (data not shown). There were significant pairwise F_{ST} values between St. Croix and Meramec, St. Croix and Clinch, and St. Croix and Tennessee populations (Table 3). The Tennessee population had significant pairwise F_{ST} values with all the other populations, and the Gasconade was also differentiated from the Clinch population, but there was no significant correlation between genetic distance and geographic distance (data not shown, $p = 0.25$). There was also no significant relationship between residuals of pairwise F_{ST} values and geographic distance (Hutchison & Templeton, 1999).

Bayesian, MP, and ML phylogenies revealed similar trees and in all of them, *C. monodonta* forms a monophyletic clade with high bootstrap and Bayesian support. However, within this clade, there is little variation, branches are short, and many haplotypes form an unresolved polytomy (Figure 4). Despite overall lack of resolution in this species, there is high posterior support for the divergence of three lineages from all other sequences, and bootstrap support over 50% for two of those lineages (Figure 4). The overall topology of the phylogenetic tree is replicated in a haplotype network (Figure 5); however, in this analysis, only the clade with both high posterior and bootstrap support forms a distinct lineage (Lineage 2, Figure 4). The bimodal mismatch distribution also indicates two lineages (Figure 3a). The proportion of lineage two was more prevalent in the St. Croix population, and decreased as latitude decreased (Figure 6), therefore, significance of this relationship was tested with linear regression ($R = 0.89$, $p = 0.057$).

Nested clade analysis of the network revealed only 3 of 10 clades had significant geographic and phylogenetic relationships sufficient to infer past events (Figure 4 and Table 4). Both clade 3-1 and its internal clade 2-2 were inferred to have restricted gene flow with isolation-by-distance (IBD), and the total cladogram inferred past gene flow with loss of intermediate populations. Mismatch distributions calculated within each lineage fit the predicted distribution under a model of sudden expansion (Figure 3b, c).

Lamarc generated a point estimate of $\theta = 0.0122$ with 95% confidence intervals of 0.0082-0.0174. Lamarc also estimated immigration rates from each population into

all others; there was an average rate of 600 (range 200-900) immigrants per generation. Based on θ, the calculated effective population size (N_e) was 18,000 (95% CI 12,000-25,000) or 19,000 (95% CI 13,000-28,000) depending on the generation time that was used (6 or 26 years, respectively). Generation times were converted to actual years based on the estimate for the divergence of lineage 2 from lineage 1 of about 850,000 (95% CI 360,000-1.6 million) years ago. Because N_e values were so close, a point estimate of N_e = 19,000 and a range (95% CI 12,500-26,500) was used in simulations of the gene tree onto population histories so that most reasonable estimates would be tested. The Bayesian phylogeny *S*-value was 32 (Slatkin & Maddison 1989), which means that 32 migration events had to occur for the haplotypes sampled in this species to be located in each population. If the minimum generation time of 6 years was simulated, the single-refugium hypothesis was accepted (indicated by arrows) for N_e = 18,000 and 26,500, but not for the lowest estimate of 12,500 (Figure 7). The dual-refugia hypothesis was rejected for divergence times of 850 thousand and 1.6 million years ago, but not for a divergence time of 360,000 years ago. The single refugium hypothesis had the most support for all but one reasonable estimate of N_e, and for the full range of divergence times; however the dual-refugium model was rejected in all but one case, the highest N_e at the youngest divergence time (data not shown).

Discussion

The lack of genetic structure in *C. monodonta* indicates that gene flow and migration have influenced its genetic diversity. Evidence of high gene flow is seen in genetic diversity indices for both allozyme and mitochondrial markers. Allozymes and mtDNA both had low θ values and low pairwise F_{ST} differences among populations. Furthermore, allozyme alleles and mtDNA haplotypes were evenly found in all distantly located populations. The COI phylogeny indicates that this species was isolated in two glacial refugia during the Pleistocene where separate lineages likely arose. Post-Pleistocene dispersal of these lineages was range-wide, further indicating high gene flow for this species. Statistical phylogeographic analyses rejected the Driftless Area as the location of the second glacial refugium.

Cumberlandia monodonta has substantial levels of genetic variation. Allozyme diversity measured by rates of polymorphism in this species ranged from 27-36%, which was not as high as in other mussel species where unionids had 5-60% polymorphism but it is much higher than other margaritiferids (European and North American) with only 0.074-0.259% polymorphism (Curole *et al*. 2004; Geist & Kuehn 2005). These populations of *C. monodonta* had much higher rates than another North American margaritiferid species, *M. hembelii,* where all 25 loci tested were fixed for a single allele (Curole *et al.* 2004). Heterozygosity values in *C. monodonta* are also higher (0.145) than values reported for margaritiferid species (0.010 – 0.018; Geist & Kuehn 2005; Machordom *et al.* 2003), but they fall within the range found in unionid species (0.0040-0.294; Curole *et al.* 2004). Differences between *C. monodonta* and other margaritiferids could be due to differences in fish hosts and stream connectivity. Margaritiferids in Europe and North America commonly use salmonids and live in parallel stream systems draining directly to the ocean. These streams lack connections to neighboring streams, preventing gene flow among drainages. This creates genetic drift leading to fixed alleles within populations and greater differentiation among populations. *Cumberlandia monodonta* is the only member of this family found in interior rivers that have connections to neighboring drainages.

Mitochondrial DNA sequences also revealed fairly high variation in *C. monodonta*. There were more haplotypes in this species compared to endangered mussels (Zanatta & Murphy 2007, 2008; Zanatta *et al.* 2007), but about the same as in other widespread species (Elderkin *et al.* 2007, 2008). Most haplotypes were only one bp different, which is why nucleotide diversity values are so low. Within-population variation was not negatively correlated with latitude, unlike the results for other mussels (Elderkin *et al.* 2008) and fish (Bernatchez & Wilson 1998). In fact the most northern population in the St. Croix had almost double the nucleotide diversity of the Meramec population in the Ozarks. Usually, the lowest nucleotide diversity values are in the north and increase as latitude decreases, presumably because southern populations were larger and more stable over time as they were not displaced by ice and neutral mutations accumulated (Avise 2000). Perhaps the St. Croix population has a higher π value than expected because it was colonized by a large contingent of mussels from

more than one stable refuge population, providing a large sample size of mtDNA haplotypes.

Both markers show that populations of *C. monodonta* are poorly differentiated and lack population structure based on low genetic distances and θ estimates that surround zero. This is in contrast to other large-river species such as *Quadrula quadrula* (Berg *et al.* 1998) and *Amblema plicata* (Elderkin *et al.* 2007) where there was significant, albeit weak, population structure. These results imply high levels of gene flow among populations, which is also supported by the lack of correlation between genetic and geographic distance. Usually, a positive correlation indicates a stepping-stone method of dispersal where the closest populations exchange genes more often than more distantly separated populations (Slatkin & Maddison 1990). For mussel species studied to date, most have met the prediction of dispersal by the stepping-stone model (Berg *et al.* 1998; Elderkin *et al.* 2007, 2008; Zanatta & Murphy 2008), which is not unexpected since mussels live in linear river systems with patchy distributions. Unfortunately, the species of fish host is unknown for *C. monodonta*, so it is difficult to describe why this species does not follow a stepping-stone model of dispersal. In most cases, mussels use fish hosts, although one species does use salamanders (Mussel Host Database, The Ohio State University Museum of Biological Diversity, http://www.128.146.250.63/bivalves). Salamanders or other unidentified hosts may be able to disperse mussels overland, by-passing nearest-neighbor populations within the river. This would be more likely in streams with high sinuosity values where large bends in the river create short overland distances compared to within-stream distances. It must be noted that very high effective population sizes can create patterns of high genetic diversity and low F_{ST} values (Waples 1998), and extant populations of *C. monodonta* do indeed number in the thousands (Baird 2000). Therefore, low levels of genetic drift could explain these results.

Even though there is no geographic population structure, this study did find genetic structure within the parsimony network and phylogenetic tree with two distinct COI lineages supported by high bootstrap values and posterior support. Amazingly, these two lineages are found across the extant range of this species covering thousands of river km. Both lineages were also present in the extirpated population

from the Tennessee River. Lineage 2 may have arisen in the north because it is more prevalent in the northern St. Croix population and decreases gradually with distance to the south, following current drainage connections in the Mississippi system.

The existence of two lineages indicates that at some point in the past, populations were isolated in two refugia, which is a common pattern found in large river mussel species studied to date (Elderkin *et al.* 2008; Hoisington *et al.* unpublished data). However, statistical phylogeographic analysis in this study rejected the two-refugium hypothesis; therefore, the Driftless Area is ruled out as a glacial refugium for this species. This hypothesis test explicitly modeled the second refuge to be in the Driftless Area; however, a refugium could have been in the headwaters of the former Teays system (Berendzen *et al.* 2003) in present day eastern OH or western WV. The ancient Teays River drained along the current valley of the upper Ohio River directly west into the upper Mississippi River, connecting eastern and midwestern streams (Figure 6). As the ice sheets spread south, they not only obliterated the Teays river basin, but cut off any movement between eastern and western drainages (Figure 6). The northern hogsucker (*Hypentelium nigricans*, Lesueur) is a fish with a range similar to that of *C. monodonta*. Lineages in this fish were estimated to have diverged pre-Pleistocene, but populations in the eastern area of its range (headwaters of the Teays River) were more closely related to populations in the upper Midwest, Iowa, and Missouri (Berendzen *et al.* 2003), reflecting river connections prior to the most recent glaciation. *Cumberlandia monodonta* was at one time found in headwater tributaries of the ancient Teays system (ORVMTS 2002), and could have been isolated there during the Pleistocene glaciations. This could also explain why lineage 2 seems to have originated in the north. Without any *C. monodonta* to sample in the old Teays system, the origin of this second lineage cannot be more closely located.

Estimates of divergence times lend support to these lineages arising during the Pleistocene followed by post-Pleistocene dispersal. Divergence was roughly estimated to occur between 36,000 to 1.6 million years ago, which falls within the Pleistocene. Post-Pleistocene dispersal of fish is thought to have occurred rapidly, aided by increased fluvial connectivity from large volumes of water as glaciers melted (Gagnon & Anders 2006). The mismatch distributions for both lineages fit a model of sudden range

expansion. Mussels would have dispersed along with fish as they moved from multiple refugia into northern habitats. Migration rates estimated in this study are in agreement with this scenario; the highest migration rate connected the two most distant populations in the Clinch and St. Croix Rivers, and was also robust between the Gasconade and Meramec populations into the St. Croix population. High migration rates are not unreasonable for this species; field studies of *M. hembeli* have shown rapid recolonization rates (10% replacement of a population in 30 months) for populations that had been completely removed by large flood events (Curole *et al*. 2004).

This study revealed that Pleistocene glaciations did drive some intraspecific diversity to arise, but post-Pleistocene dispersal had a stronger affect on the structuring of population genetics in *C. monodonta*. Both nuclear allozyme markers and mtDNA sequences provide evidence of high gene flow post-Pleistocene because all of the populations in this study are genetically similar. Over time, these populations will likely become genetically structured for two reasons — first, because the rivers connecting these distant populations are impounded and secondly, the loss of intermediate populations prevents any gene flow via the stepping stone model of dispersal. Furthermore, because this species has been extirpated from most of its former range, there is not much opportunity to understand the evolutionary history of this species, unless fish host(s) are determined. Then comparative phylogeography between host fish and *C. monodonta* may help locate where isolation occurred during the Pleistocene. The identity of the fish host may help us to understand how this mussel enjoyed such high levels of gene flow prior to its decline.

References

Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, Massachusetts.

Baird MS (2000) Life history of the spectaclecase, *Cumberlandia monodonta* Say, 1829 (Bivalvia, Unionoidea, Margaritiferidae). MS, Missouri State University.

Berendzen PB, Simons AM, Wood RM (2003) Phylogeography of the northern hogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. *Journal of Biogeography*, **30**, 1139-1152.

Berg DJ, Haag WR, Guttman SI, Sickel JB (1995) Mantle biopsy: a technique for non-destructive tissue sampling of freshwater mussels. *Journal of the North American Benthological Society*, **14**, 577-581.

Berg DJ, Cantonwine EG, Hoeh WR, Guttman SI (1998) Genetic structure of *Quadrula quadrula* (Bivalvia: Unionidae): little variation across large distances. *Journal of Shellfish Research*, **17**, 1365-1373.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Neararctic and Palearctic fish. *Molecular Ecology*, **7**, 431-452.

Buhay JE, Serb JM, Dean CR, Parham Q, Lydeard C (2002) Conservation genetics of two endangered unionid bivalve species, *Epioblasma florentina walkeri* and *E. capsaeformis* (Unionidae: Lampsilini). *Journal of Molluscan Studies*, **68**, 385-391.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657-1660.

Curole JP, Foltz DW, Brown KM (2004) Extensive allozyme monomorphism in a threatened species of freshwater mussel, *Margaritifera hembeli* Conrad (Bivalvia: Margaritiferidae). *Conservation Genetics*, **5**, 271-278.

Elderkin CE, Christian AD, Vaughn CC, Metcalfe-Smith JL, Berg DJ (2007) Population genetics of the freshwater mussel, *Amblema plicata* (Say 1817)(Bivalvia: Unionidae): evidence of high dispersal and post-glacial colonization. *Conservation Genetics*, **8**, 355-372.

Elderkin CE, Christian AD, Metcalfe-Smith JL, Berg DJ (2008) Population genetics and phylogeography of freshwater mussels in North America, *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia: Unionidae). *Molecular Ecology*, **17**, 2149-2163.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783-791.

Fetzner JW, Crandall KA (2003) Linear habitats and the nested clade analysis: an empirical evaluation of geographic versus river distances using an Ozark crayfish (Decapoda: Cambaridae). *Evolution*, **57**, 2101-2118.

Gagnon MC, Angers B (2006) The determinant role of temporary proglacial drainages on the genetic structure of fish. *Molecular Ecology*, **15**, 1051-1065.

Geist J, Kuehn R (2005) Genetic diversity and differentiation of central European freshwater pearl mussel (*Margaritifera margaritifera* L.) populations: implications for conservation and management. *Molecular Ecology*, **14**, 425-439.

Gibbard P, Van Kolfschoten T (2004) The Pleistocene and Holocene Epochs. In *A Geologic Time Scale* (eds Gradstein FM, Ogg JG, Smith AG), pp. 441-452. Cambridge University Press, Cambridge.

Graf DL (1997) Northern redistribution of freshwater pearly mussels (Bivalvia: Unionoidea) during Wisconsin deglaciation in the southern glacial Lake Agassiz region: a review. *American Midland Naturalist*, **138**, 37-47.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696-704.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.

Hebert PDN, Beaton MJ (1993) *Methodologies for allozyme analysis using cellulose acetate electrophoresis*. Unpublished report. Helena Laboratories, Beaumont, Texas.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnaean Society,* **58**, 3, 247-276.

Hove MC, Kurth JE, Kapuscinski AR (1998) Brown bullhead suitable host for *Tritogonia verrucosa*; *Cumberlandia monodonta* host(s) remain elusive. Triannual Unionid Report, Issue 15, 13. [Online] http://ellipse.inhs.uiuc.edu/FMCS/TUR/TUR15.html [accessed 11 May 2008].

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754-755.

Huff SW, Campbell D, Gustafson DL, et al. (2004) Investigations into the phylogenetics of freshwater pearl mussels (Bivavlvia: Margaritiferidae) based on molecular data: implications for their taxonomy and biogeography. *Journal of Molluscan Studies*, **70**, 379-388.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, **53**, 1898-1914.

King TL, Eackles MS, Gjetvaj B, Hoeh WR (1999) Intraspecific phylogeography of *Lasmigona subviridis* (Bivalvia: Unionidae): conservation implications of range discontinuity. *Molecular Ecology*, **8**, S65-S78.

Knowles LL (2004) The burgeoning field of statistical phylogeography. *Journal of Evolutionary Biology*, **17**, 1-10.

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, **22**, 768-770.

Kuhner MK, Smith LP (2007) Comparing likelihood and Bayesian coalescent estimation of population parameters. *Genetics*, **175**, 155-165.

Lydeard C, Cowie RH, Ponder WF et al. (2004) The global decline of nonmarine molluscs. *Bioscience*, **54**, 321-330.

Machordom AR, Araujo R, Erpenbeck D, Ramos MA (2003) Phylogeography and conservation genetics of endangered European Margaritiferidae (Bivalvia: Unionoidea). *Biological Journal of the Linnaean Society*, **78**, 235-252.

Maddison WP (2007) Coalescence package for mesquite. Version 2.01. http://mesquiteproject.org.

Maddison WP, Maddison DR (2007) Mesquite: a modular system for evolutionary analysis. Version 2.01 http://mesquiteproject.org.

Mandrak NE, Crossman EJ (1992) Postglacial dispersal of freshwater fish into Ontario. *Canadian Journal of Zoology*, **70**, 2247-2259.

Marko PB (2005) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Molecular Biology and Evolution*, **19**, 2005-2021.

Miller MP (1997) Tools for population genetics analyses, TFPGA. A Windows program for the analysis of allozyme and molecular population genetic data. Available from the author: MP Miller http://www.bioweb.usu.edu/mpmbio/.

Near TJ, Page LM, Mayden RL (2001) Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the central highlands pre-Pleistocene vicariance hypothesis. *Molecular Ecology*, **10**, 2235-2240.

Nei M (1978) Estimating average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583-590.

Nekola, JC (1999) Paleorefugia and neorefugia: the influence of colonization history on community pattern and process. *Ecology*, **80**, 2459-2473.

Nilsson C, Reidy CA, Dynesius M, Revenga C (2005) Fragmentation and flow regulation of the world's large river systems. *Science*, **308**, 405-408.

Nylander JA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Ohio River Valley Ecosystem Team Mollusk Subgroup (2002) Status assessment report for the spectaclecase, *Cumberlandia monodonta*, occurring in the Mississippi River system (U.S. Fish and Wildlife Service Regions 3, 4, 5, and 6). 1 Sept, pp. 1-6. Online] at: www.fws.gov/Midwest/endangered/clams/spectaclecase-sa.pdf [accessed 12 May 2008]

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487-488.

Posada D, Crandall KA, Templeton AR (2006) Nested clade analysis statistics. *Molecular Ecology Notes*, **6**, 590-593.

Richards CL, Carstens BC, Knowles LL (2007) Distribution modeling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, **34**, 1833-1845.

Richardson BJ, Baverstock PR, Adams M (1986) *Allozyme electrophoresis: a handbook for animal and population studies*. Academic Press, New York.

Roe KJ, Hartfield PD, Lydeard C (2001) Phylogeographic analysis of the threatened and endangered superconglutinate-producing mussels of the genus *Lampsilis* (Bivalvia: Unionidae). *Molecular Ecology*, **10**, 2225-2324.

Ross TK (1999) Phylogeography and conservation genetics of the Iowa Pleistocene snail. *Molecular Ecology*, **8**, 1363-1373.

Roy K, Valentine JW, Jablonski D, Kidwell SM (1996) Scales of climatic variability and time averaging in Pleistocene biotas: implications for ecology and evolution. *Trends in Ecology and Evolution*, **11**, 458-463.

Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning, a laboratory manual*. 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Sei M, Porter AH (2007) Delimiting species boundaries and the conservation genetics of the endangered maritime ringlet butterfly (*Coenonympha nipisiquit* McDunnough). *Molecular Ecology*, **16**, 3313-3325.

Slatkin M, Maddison, WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603-613.

Slatkin M, Maddison WP (1990) Detecting isolation by distance using phylogenies of genes. *Genetics*, **126**, 249-260.

Smith I (2007) Historical biogeography: the new synthesis. *Current Biology*, **17**, 598-600.

Strange RM, Burr BM (1997) Intraspecific phylogeography of North American highland fish: a test of the Pleistocene vicariance hypothesis. *Evolution*, **51**, 885-897.

Swofford DL, Selander RB (1981) BIOSYS-1: a FORTRAN program for the comparative analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity*, **72**, 281-283.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology Evolution*, **24**, 1596-1599.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381-397.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789-809.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping.1. basic theory and an analysis of alcohol-dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343-351.

Templeton AR, Sing CF (1993) A cladistic-analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. 4. nested analyses with cladogram uncertainty and recombination, *Genetics*, **134**, 659-669

Turner TF, Trexler JC, Harris JL, Haynes JL (2000) Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. *Genetics*, **154**, 777-785.

Vucetich JA, Waite TA (2003) Sspatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. *Conservation Genetics*, **4**, 639-645.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438-450.

Wares JP, Turner TF (2003) Phylogeography and diversification in aquatic mollusks. In: *Molecular Systematics and Phylogeography of Mollusks* (eds Lydeard C, Lindberg DR), pp. 229-269. Smithsonian Books, Washington D.C.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Whiteley AR, Spruell P, Allendorf FW (2004) Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. *Molecular Ecology*, **13**, 3675-3688.

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, **19**, 395-420.

Wright S (1978) *Evolution and the genetics of populations, volume 4, variation within and among natural populations*. University of Chicago Press, Chicago, Illinois.

Zanatta DT, Murphy RW (2007) Range-wide population genetic analysis of the endangered northern riffleshell mussel, *Epioblasma torulosa rangiana* (Bivalvia: Unionoida). *Conservation Genetics*, **8**, 1393-1404.

Zanatta DT, Fraley SJ Murphy RW (2007) Population structure and mantle display polymorphisms in the wavy-rayed lampmussel, *Lampsilis fasciola* (Bivalvia: Unionidae). *Canadian Journal of Zoology,***85**, 1169-1181.

Zanatta DT, Murphy RW (2008) The phylogeographical and management implications of genetic population structure in the imperilled snuffbox mussel, *Epioblasma triquietra* (Bivalvia: Unionidae). *Biological Journal of the Linnaean Society*, **93**, 371-384.

Zouros E, Ball AO, Saavedra C, Freeman KR (1994) Mitochondrial DNA inheritance. *Nature*, **368**, 818.

Table 1. Descriptive statistics of mtDNA data for each collection site of *Cumberlandia monodonta* from populations in the Clinch River, TN; Gasconade and Meramec Rivers, MO; St. Croix River, MN; Tennessee River, AL; and a single specimen in the Cumberland River, TN. Note: Tennessee River populations were combined into one for analyses.

River	n	Total number of haplotypes	Number of unique haplotypes	Mean number of base-pair differences	Nucleotide diversity (π)
Clinch	48	18	11	4.17	0.0062
Gasconade	40	9	6	3.18	0.0047
Meramec	34	8	3	3.03	0.0015
St. Croix	48	9	4	2.11	0.0031
TN 1	2	6	1	4.86	0.0073
TN 2	5				
Cumberland	1	1	0	-	-

Table 2. Allele frequencies at polymorphic allozyme loci for *Cumberlandia monodonta* from populations in the Clinch River, TN; Gasconade River, MO; and St. Croix River, MN.
N = number of individuals.

		Population		
Locus	Allele	Clinch	Gasconade	St. Croix
PGM-Fast		N = 49	N = 43	N = 50
	1	0.551	0.616	0.570
	2	0.367	0.314	0.370
	3	0.031	0.070	0.040
	4	0.051	0.000	0.020
PGM-Slow			N = 44	
	1	-	0.080	-
	2	-	0.920	-
	3	-	0.000	-
LAP		N = 48	N = 42	N = 50
	1	0.396	0.357	0.330
	2	0.604	0.631	0.660
	3	0.000	0.000	0.010
	4	0.000	0.012	0.000
PEP		N = 49	N = 42	N = 50
	1	0.388	0.345	0.310
	2	0.612	0.643	0.680
	3	0.000	0.000	0.010
	4	0.000	0.012	0.000
Mean (± SE) sample size per locus		46.7 (1.7)	41.4 (1.3)	48.6 (1.0)
Mean (± SE) number of alleles per locus		1.7 (0.3)	1.8 (0.3)	2.2 (0.3)
Percent polymorphic loci (95% criterion)		27.3	36.4	27.3
Mean heterozygosity		0.142	0.133	0.159

Table 3. Pairwise F_{ST} values for cytochrome oxidase I sequences from five populations of *Cumberlandia monodonta*.

Population	Gasconade	St. Croix	Meramec	Clinch	Tennessee
Gasconade	0.000	-	-	-	-
St. Croix	0.011	0.000	-	-	-
Meramec	-0.003	0.041*	0.000	-	-
Clinch	0.041*	0.129*	0.015	0.000	-
Tennessee	0.115*	0.171*	0.109*	0.125*	0.000

Estimates were tested for significance* at $p = 0.05$ with 1023 nonparametric permutations of the data.

Table 4. Nested clade analysis of geographical associations for populations of *Cumberlandia monodonta*.

Clade	Permutational chi-square statistic	Probability	Inference Key # and outcome
2-2	24.4	0.03	1,2,3,4 restricted gene flow with IBD
3-1	34.0	0.00	1,2,3,4 restricted gene flow with IBD
Total cladogram	132.2	0.00	1,2,3,5,6,7,8 past gene flow then loss of intermediate populations

The clade column refers to nesting clades in Figure 6, and clades with neither genetic nor geographic variation are not included because no test is possible. IBD = isolation-by-distance.

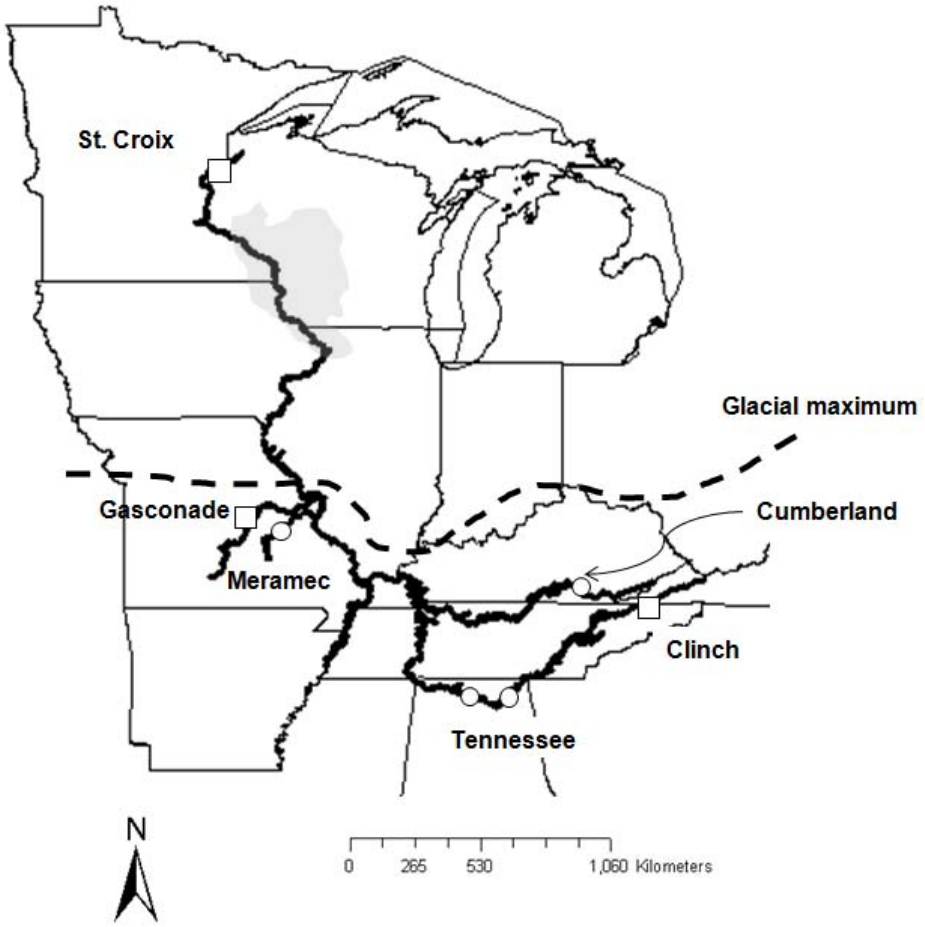

Figure 1. Map of the central United States indicating sites where *Cumberlandia monodonta* were sampled. Populations are labeled with the name of their resident rivers. All sites represent populations, except the Cumberland which was a singleton. Square symbols are populations with allozyme and mtDNA analyses, circles are populations analyzed for mtDNA only. Driftless Area indicated by shaded region, dashed line is the glacial maximum.

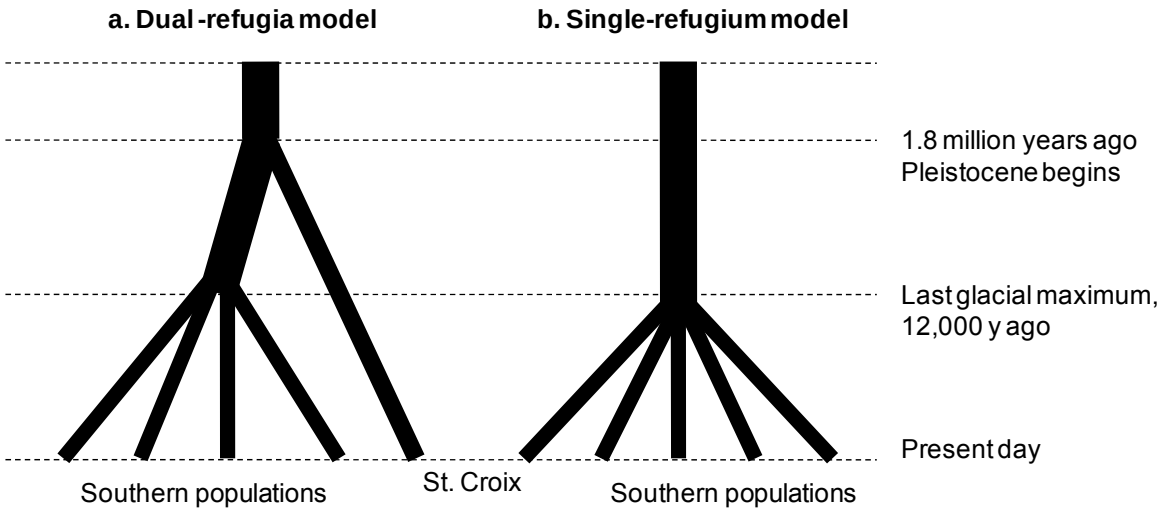

Figure 2. Models of geographic configurations for populations of *Cumberlandia monodonta* used to determine if genetic variation reflects colonization from (a) dual-refugia including the Driftless Area or (b) a single-refugium.

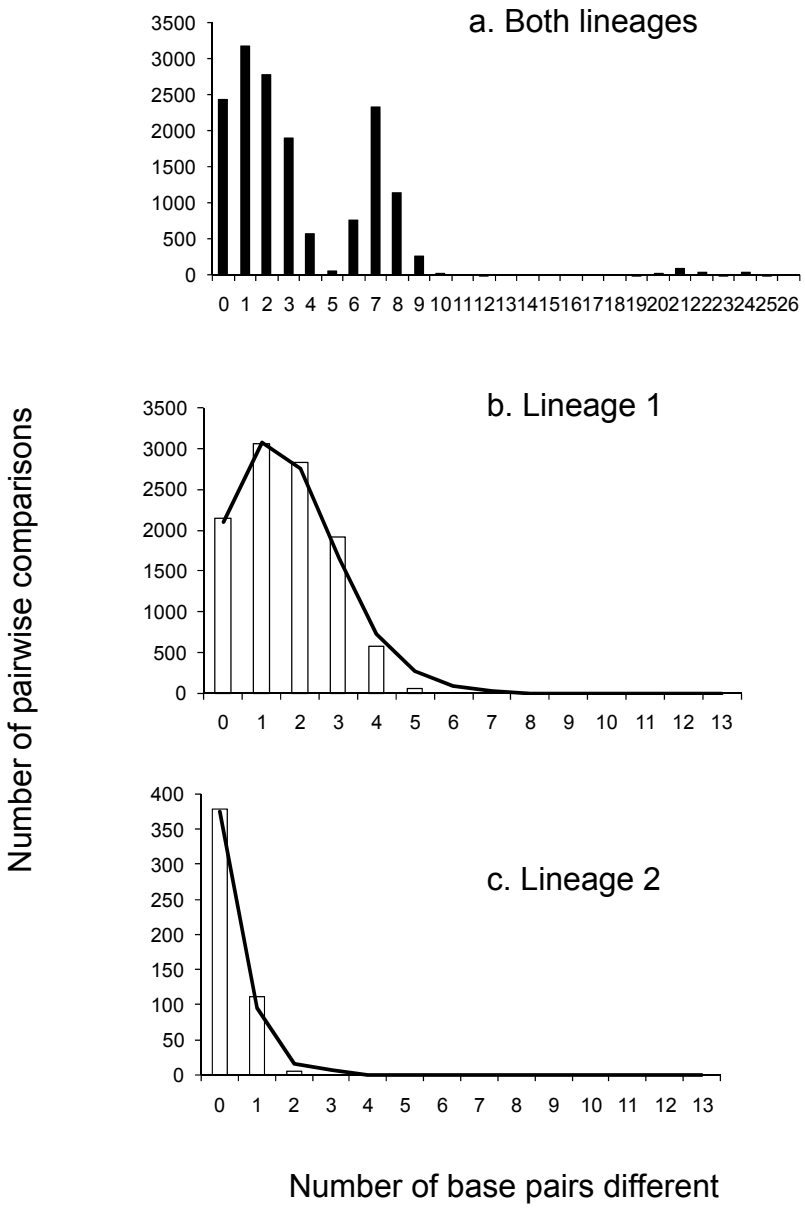

Figure 3. Mismatch distributions over all sequences and for each lineage. Frequency of the number of base pairs that differ at the COI locus between each pair of *Cumberlandia monodonta* individuals, a) over all samples, b) individuals within lineage 1, c) individuals within lineage 2. Bars in lineage charts are observed data; lines are predicted distribution under a model of sudden expansion for each lineage.

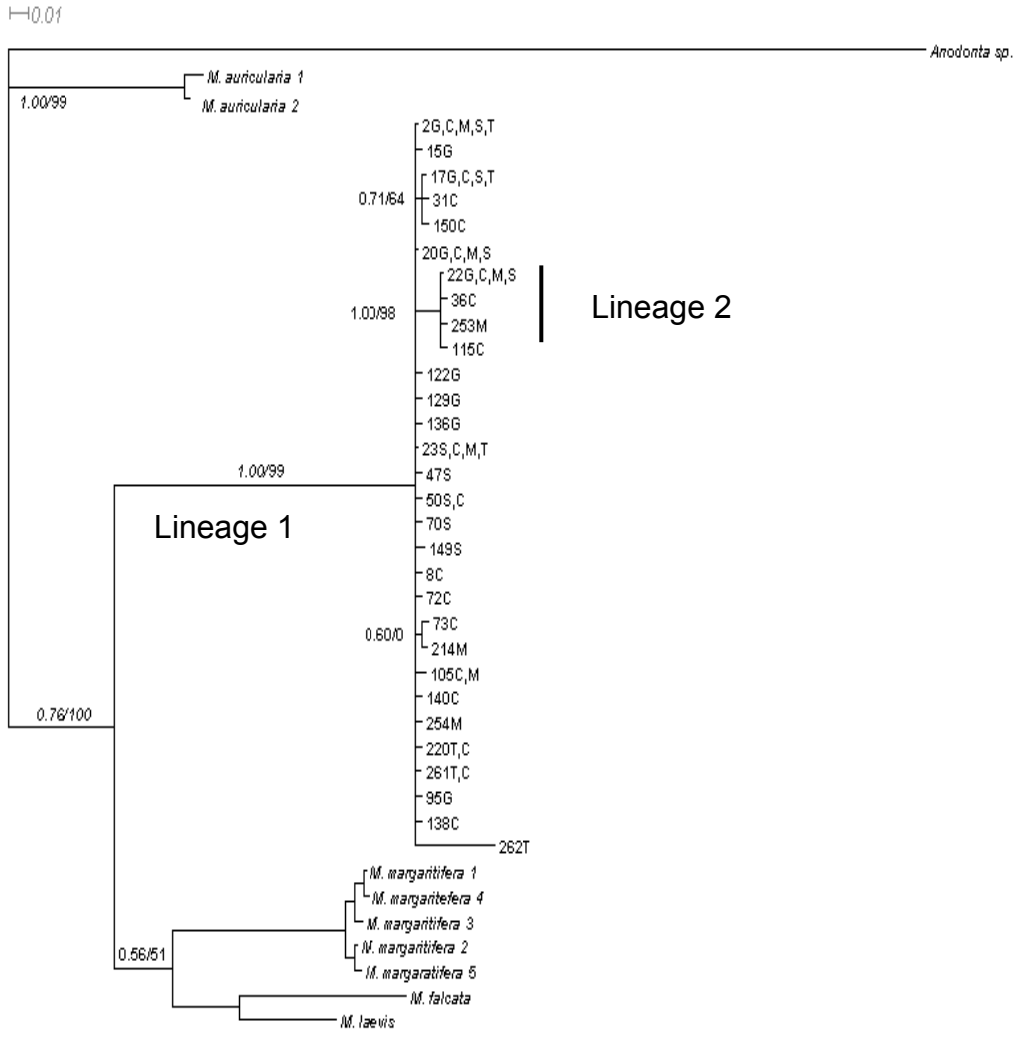

Figure 4. Bayesian inference of phylogeny for Margaritiferidae. *Cumberlandia monodonta* haplotypes are numbered according to sample ID, letters designate initial of populations where they are located as labeled in Figure 1. Decimal values of support at the nodes are Bayesian posterior probabilities from 15,800 trees, and bootstrap values greater than 50% are reported second for maximum parsimony analyses. Sequences from the genus *Margaritifera* are from GenBank, as is the outgroup sequence.

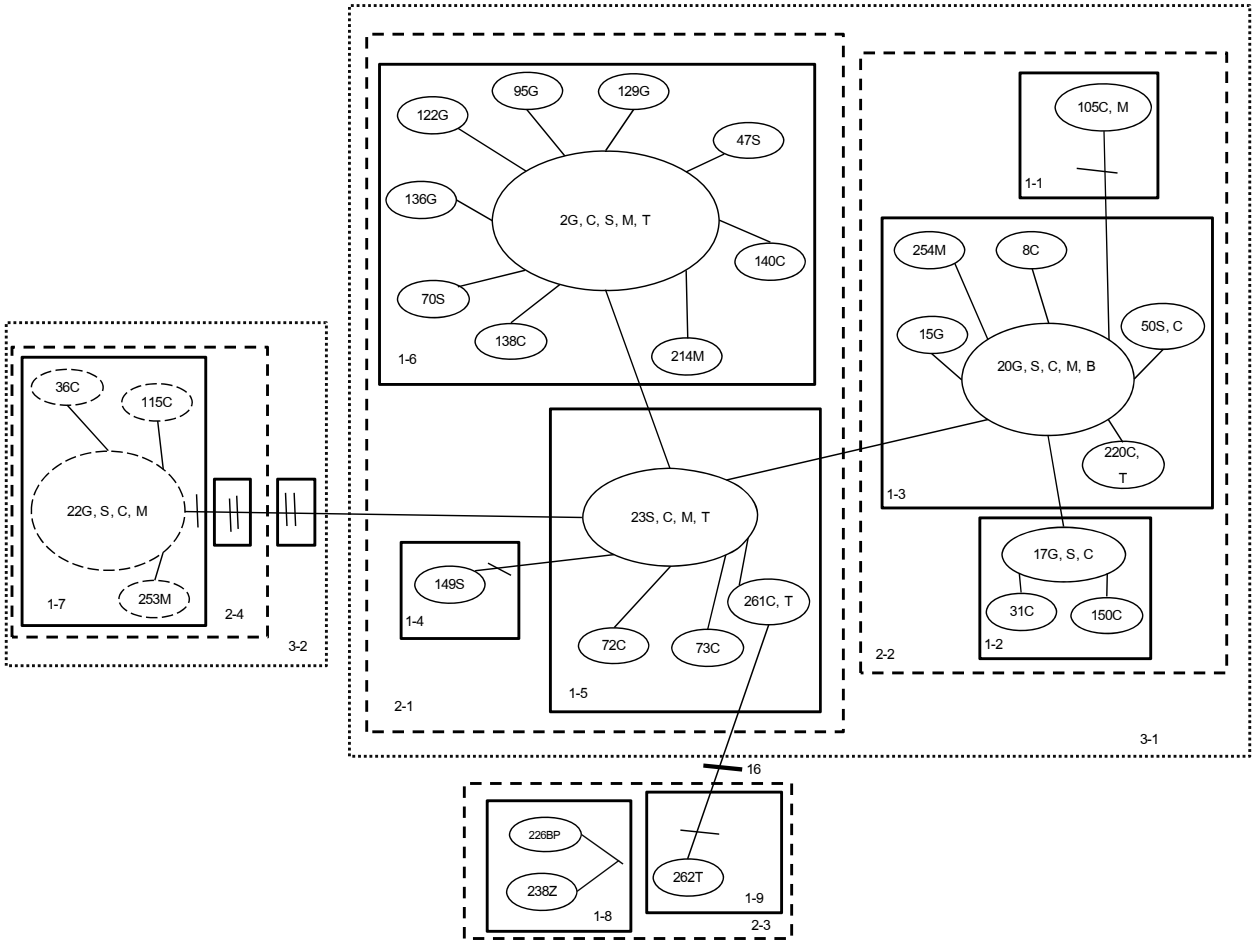

Figure 5. Parsimony network with nested clades for *Cumberlandia monodonta* showing 1-, 2-, and 3-step clades that contain genetic and geographical variation (letters indicate locations by first initial as in Figure 1; B = Cumberland River). Haplotypes are numbered as in Figure 4, each line represents one base-pair difference between haplotypes, hashes are inferred missing haplotypes (note bold hash indicates more than one base-pair difference), and haplotype frequency is relative to the size of the ellipse. Lineage 2 is contained within clade 3-2.

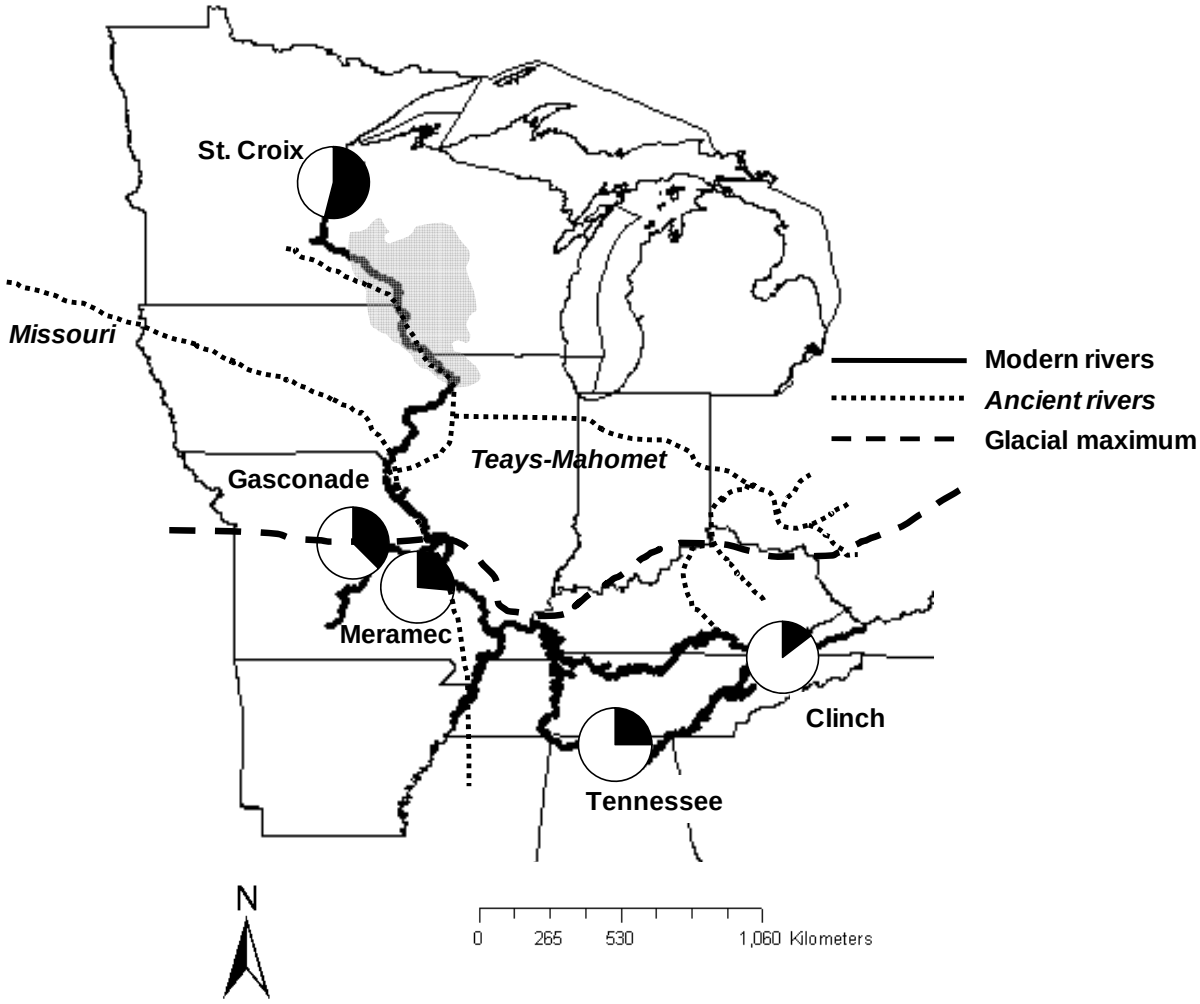

Figure 6. Map of the central United States indicating current and ancient rivers with sites where populations of *Cumberlandia monodonta* were sampled. Populations are labeled as in Figure 1; pie charts show the frequency of linage 1 (■) and lineage 2 (□) resolved in the cytochrome oxidase I phylogeny. Driftless Area is indicated by shaded area. Dashed line is the glacial maximum.

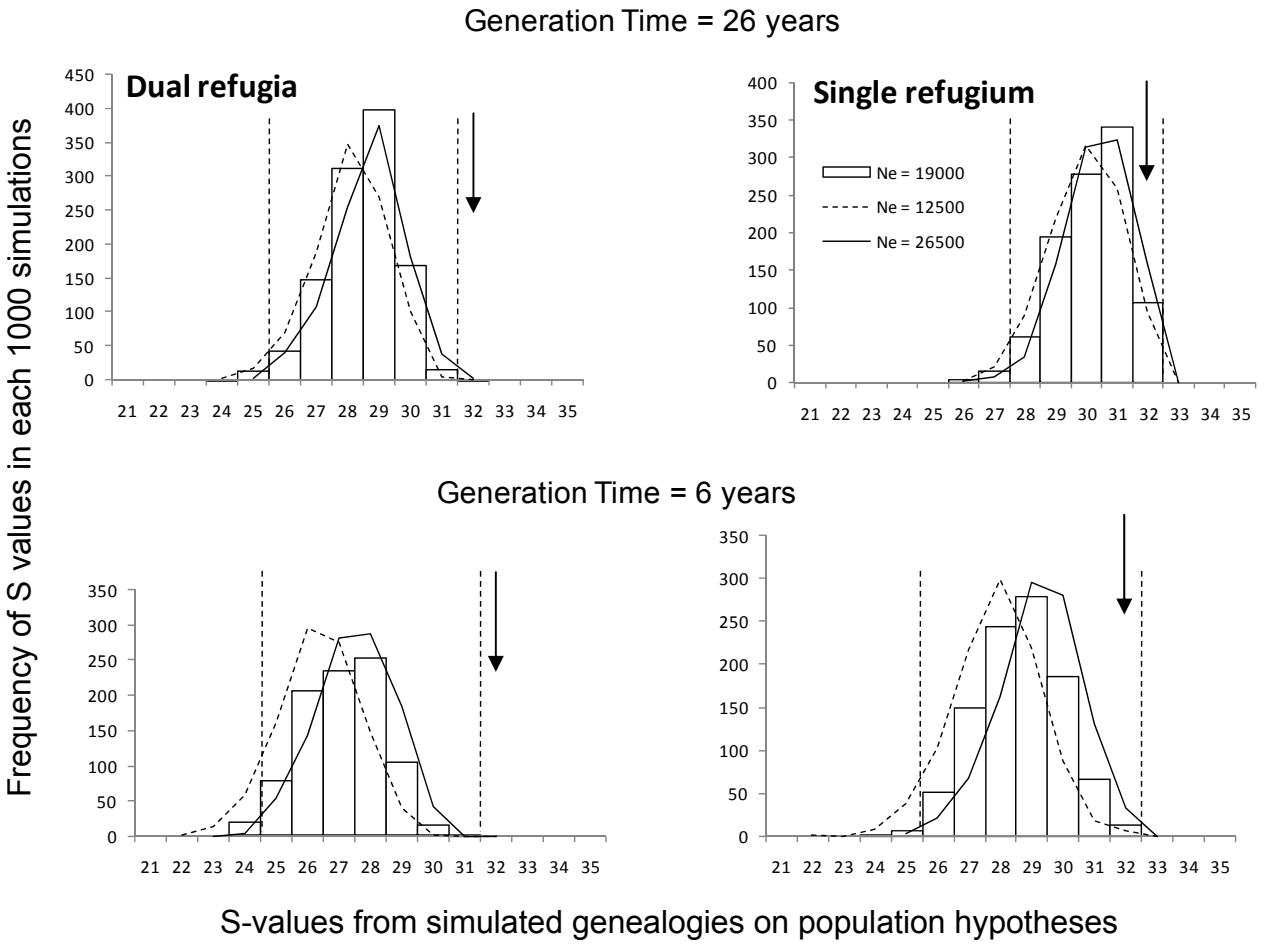

Figure 7. Frequency histograms of S-values for simulated genealogies onto dual-refugia and single-refugium population hypotheses, for two generation times (top 26 years, bottom 6 years). Bar charts are coalesced with the point estimate of N_e, and lines are for the minimum and maximum N_e estimates. Vertical dotted lines indicate the 95% confidence intervals for distributions of bar charts, S-values of 32 indicate acceptance (arrows) of the model.

Chapter 3

Phylogeography of *Lampsilis cardium*: genetic patterns reflect post-Pleistocene dispersal

Abstract

Pleistocene glaciations have created genetic structure in populations of North American aquatic fauna. Common patterns in genetic diversity from glacial events include gradients of genetic diversity that increase from northern to southern areas and evidence of multiple glacial refugia within the phylogeny of mitochondrial genes. Two biogeographic hypotheses for colonization of the Red River of the North (RRN) have been generated by fish and mussel distributions. One invokes post-Pleistocene dispersal from the upper Mississippi River (UM) into the RRN drainage system and the other suggests that dispersal occurred into the RRN and Lake of the Woods drainages via glacial Lake Agassiz at the end of the Pleistocene. COI gene sequences were used to evaluate population genetic structure in *Lampsilis cardium*, a mussel that is widespread across the Mississippi drainage system. Phylogenetic and geographic data were used to determine which dispersal process into the RRN was most likely. This species has high levels of genetic diversity but weak population structure from Arkansas to Canada, which is likely due to high gene. Genetic diversity indices were not significantly associated with latitude. Shallow divergence within the COI gene tree indicated isolation during the Pleistocene in at least two refugia for the population from the Great Lakes drainage, a pattern which supports findings for several other mussel species. Genetic data from the UM and tributaries of the RRN did not refute northern dispersal to several drainage systems via Lake Agassiz.

Introduction

Species distributions and population genetic structure are determined by extrinsic (i.e. landscape history) and intrinsic factors (i.e. life-history characteristics; (Vucetich & Waite 2003; Wares & Turner 2003; Whiteley *et al.* 2004). Pleistocene glaciations had profound effects on the distribution and population genetics of many species (Hewitt 1996). Freshwater systems were the most altered ecosystems during the ebb and flow of several glacial events which altered drainage patterns as ice sheets surged and receded (Bernatchez & Wilson 1998). Extensive work on the distribution and biogeography of fish has (Berendzen *et al.* 2003; Bernatchez & Wilson 1998; Hardy *et al.* 2002; Near *et al.* 2001; Ray *et al.* 2006; Switzer 2004) examined the effects of glaciations on species distributions and population genetic structure, and determined recolonization routes from glacial refugia (Mandrak & Crossman 1992). During periods of high water due to glacial melting, large areas of land were inundated, allowing for wide dispersal of aquatic organisms and facilitating mixture of intraspecific populations from different refugia (Bernatchez & Wilson 1998; Gagnon & Angers 2006). The combination of varying dispersal routes before and after the glaciers has lead to different patterns in population genetic structure for different species (Strange & Burr 1997).

Comparative phylogeographic analyses on fish species with ranges in previously glaciated (northern) regions and regions that were never glaciated (southern) reveal consistent patterns (Bernatchez 2001). Northern populations have lower levels of genetic diversity compared to southern populations along with less intraspecific divergence (Bernatchez 2001). Colonization of northern areas 8,000-12,000 years ago is often evident in genetic data that indicates range expansion and rapid population growth (Bernatchez 2001). Furthermore, populations in the north have been colonized from multiple glacial refugia (Bernatchez 2001). Since fish are obligatory hosts during part of the freshwater mussel life cycle, mussels were likely carried and dispersed into formerly glaciated areas by fish, and therefore mussels and their host fish should have similar patterns in distribution and genetic structure.

There are about 300 freshwater mussel species in North America, and several species have distribution ranges from Canada to the Gulf of Mexico (Cummings & Mayer 1992; Parmalee & Bogan 1998). Host fish, necessary to transform mussel larvae

to juveniles, influence the downstream and upstream distribution of mussels. Mussels that use migratory fish are likely to have less population genetic differentiation than mussels that use fish with small home ranges. Reproductive patterns vary among mussel species, with some larvae parasitizing fish during the spring which may promote dispersal due to high water, while others remain on fish in the fall when low water levels may limit fish movement. Large-river mussel species have very little genetic population structure over great distances, and populations north of the glacial maximum often have lower levels of genetic diversity than their counterparts in the south (Elderkin *et al.* 2007, 2008; Zanatta & Murphy 2008). Even in species that occupy medium to small streams, populations in the north have less diversity than those in the south (Zanatta & Murphy 2008). Biogeographic studies have shown that mussels colonized new habitats from multiple refugia (Graf 1997, 2002), which is concordant with fish (Mandrak & Crossman 1992). Genetic diversity of populations in the north has confirmed colonization from multiple refugia for some mussel species (Elderkin *et al.* 2007, 2008), but not for others (Zanatta & Murphy 2007, 2008).

The Canadian Interior, Upper Mississippi (UM), and Great Lake basins were heavily affected by the last Wisconsin ice sheet of the Pleistocene era (Graf 1997). Glacial Lake Agassiz connected upper midwest drainage systems of the Great Lakes, UM, Red River of the North (RRN), and Lake of the Woods (LOW; RRN and LOW are now part of the Nelson drainage basin into Hudson Bay), and as ice melted, it was drained by two outlets. One was Glacial River Warren, whose riverbed is now occupied by the RRN and Minnesota Rivers, and the second followed the current valley of the St. Croix River. Lake Agassiz, which once covered all of the current drainage systems in the area, may have been the agent for mussel dispersal into all of these drainages. If this is true, then fish and mussel species should be shared within these systems. However, this is not the case (Graf 1997). Fauna in the LOW, UM (Lake Itasca to Minneapolis and the St. Croix rivers), and Lake Superior drainages are not shared with the RNN system; only the UM and RRN drainages have the same fauna (Graf 1997; Radke 1992). An alternate explanation for dispersal of UM fauna into the RRN is *post*-Lake Agassiz dispersal via stream head capture. Radke (1992) proposed that the Otter Tail River of western MN was captured by the RRN *after* Lake Agassiz was drained (8600 years ago). This would explain common species between the UM and RRN, and the lack of species shared between the RRN, LOW, and Lake Superior drainages (Graf

1997; Radke 1992). *Post*-glacial dispersal is in contrast to several other studies which show proglacial lakes, such as Lake Agassiz, were often used to disperse fauna far and wide (Bernatchez 2001; Gagnon & Angers 2006).

Lampsilis cardium is widely distributed in the entire Mississippi, RRN, and LOW systems in previously glaciated areas; and in southern areas that were never glaciated. I evaluated patterns of genetic diversity between populations in the north versus the south to determine if this large-river species has similar patterns found in other species of fish and mussels. Populations in the Upper Midwest and Canadian drainages were used to determine if stream head capture of the Otter Tail River allowed post-Pleistocene dispersal from the UM into the RRN. Based on previous studies of fish and mussels, there are some explicit hypotheses to test: 1) Genetic diversity in populations in previously glaciated regions is lower than in populations from southern areas that were non-glaciated. This predicts that there will be fewer haplotypes and lower nucleotide diversity in northern populations relative to southern populations. 2) Populations in the north were colonized from multiple refugia, thus northern populations should have genetic evidence of range expansion. 3) If the Otter Tail River was captured post-Pleistocene by the RRN, then populations in the Pomme de Terre River should be more similar to populations in the RRN system than they are to populations in the LOW system. This predicts lower pairwise F_{ST} values, indicating higher similarity, between Pomme de Terre and RRN populations than between RRN and LOW populations.

Materials and methods

Sample collection

Samples were obtained from 13 populations in six drainages (Figure 1). One population was obtained from Lake St. Clair (Great Lakes), the Rainy River (LOW), Pomme de Terre River (tributary of the Minnesota River; UM), and the Boone River (UM). Two populations were sampled in the Assiniboine system and one in the Roseau River, all part of the RRN system. Two populations from the UM were sampled. These nine populations are all north of the glacial maximum. Another four populations were sampled south of the glacial maximum: two in the lower Mississippi drainage in the Spring and Ouachita Rivers in AR; and two in the lower Ohio system in the Green and Licking Rivers of KY (Figure 1). Tissue clips were taken from the mantle of 116 animals

in the field and preserved in 95% ethanol or on dry ice and frozen at -80 ^{o}C until DNA was extracted.

Analyses

Total genomic DNA was extracted from tissues using ~25 mg tissue and DNeasy (Qiagen Inc., Valencia, CA) animal tissue kits. Polymerase chain reaction (PCR) amplification of 633 base-pairs of the COI gene was performed using primers from Walker *et al.* (2006); LCO22me2 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO700dy2 (5'-TCAGGGTGACCAAAAAATCA- 3') (Midland Certified Reagent Company, Midland, TX). In a final volume of 50 µl, the PCR mix contained 0.01 µg genomic DNA, 0.5 µM each primer, 4.0 mM $MgCl_2$, and *Taq* master mix (Qiagen Inc.) that provided 200 µM of each dNTP, and 2.5 U *Taq*. Amplification was performed in PTC-200 or PTC-100 thermocyclers (MJ Research, Waltham, MA) with an initial denaturation of 94ºC for 120 seconds followed by 35 cycles of 30 seconds at 94ºC, 30 seconds at 40ºC, 90 seconds at 72ºC, and an extension phase at 72ºC for 210 seconds. PCR products were run on 2% agarose gels and gel-isolated using the QIAquick Spin Kit (Qiagen Inc.). To ensure accurate base calls, DNA samples were cycle sequenced in both directions with BigDye v3.1 (Applied Biosystems Inc. Foster City, CA, ABI) and analyzed on an ABI Prism 3100 or 3730 sequencer (ABI). Sequences were aligned by eye, using only base-pair (bp) substitutions in BioEdit 7 (Hall 1999); all polymorphic sites were confirmed with the raw electropherograms.

Data analysis

Genetic diversity was quantified as the number of haplotypes, the mean number of different bp, and nucleotide diversity (π; Nei 1987) in Arlequin 3.1 (Schneider *et al.* 2000). The number of haplotypes and nucleotide diversity within populations was tested for correlation with latitude using linear regression analysis. A phylogeny was estimated with maximum parsimony (MP) in MEGA 4.0 (Tamura *et al.* 2007). Three *Lampsilis ovata* sequences, two from the Licking River site and a third from Genbank (AY613826), along with a single *Cumberlandia monodonta* sequence were used to root the tree. Tree topology was evaluated with 1000 bootstrap replicates, and branches with over 50% support were considered strong. Maximum parsimony was also used to construct a network of all individual gene sequences using the program TCS version 1.13 (Clement *et al.* 2000). This analysis depicts relationships among haplotypes in a

network and it also displays the frequency of haplotypes. Multiple connections or loops among haplotypes were simplified using a coalescent approach (Fetzner & Crandall 2003), where mutations are more probable if they arose from a more frequent haplotype.

Genetic differentiation was analyzed within drainages with more than one population sampled and among drainages over a large geographic area. Populations were split into six groups: Great Lakes, Rainy River, RRN (Assiniboine, Souris, and Roseau Rivers), UM (including Pomme de Terre and Boone Rivers), Ohio River (KY sites), and lower Mississippi River (AR sites). Differentiation was quantified with analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) and significance was tested using 1023 nonparametric permutations with $\alpha = 0.05$ (Weir & Cockerham 1984). Further evidence of genetic population structure was quantified by calculating F_{ST} values (Tamura & Nei 1993) for each pair of populations using 1023 permutations and significance determined at $\alpha = 0.05$. Isolation-by-distance (IBD) was tested for using a Mantel test for matrices of pairwise F_{ST} values and geographic distances measured by linear river distances (km). Geographic distances between each pair of populations were measured to the nearest km using the path measurement tool in GoogleEarth© (www.googleearth.com) and following each bend in the river. Populations were placed into GoogleEarth with coordinates obtained onsite with a Garmin (Olathe, KS) handheld Global Positioning Unit. Drainage systems were measured through temporary connections established as glacial lakes drained at the end of the Pleistocene, which are described below. Evidence of demographic expansion was evaluated with mismatch distributions calculated in Arlequin where a Poisson distribution would be expected if populations had recently expanded (Rogers & Harpending 1992; Slatkin & Hudson 1991). Mismatch distributions were calculated with 1000 replicates for each population and for populations grouped as northern or southern (Table 1).

To evaluate the hypothesis of a post-Pleistocene connection between the UM and the RRN, the relationship among populations in the Pomme de Terre and Rainy rivers and tributaries of the RRN were evaluated. A haplotype network was constructed in TCS with maximum parsimony for only those sequences found in these three systems. Pairwise F_{ST} values were used to determine differences among populations. Pomme de Terre and RRN populations should have lower pairwise F_{ST} values than Pomme de Terre and Rainy River populations.

Further exploration of historical effects of the Pleistocene was conducted with nested clade analysis (NCA; Templeton *et al.* 1995; Templeton 1998; Templeton 2004). Nested clade analysis is a powerful methodology that adds a temporal perspective to spatial population genetics (Fetzner & Crandall 2003). Templeton's work, and that of many others, has used NCA for terrestrial species where geographic distances are measured as the shortest distance between the centers of each population. In contrast river distances have been used in NCA for aquatic species (Bernatchez 2001; Fetzner & Crandall 2003) because this measure reflects the actual distance freshwater animals must travel between populations. Therefore, NCA was conducted on the haplotype network using the program Geodis (Posada *et al.* 2000) with linear river distances (km).

Both NCA and Mantel tests for IBD required river distances to be measured between currently disconnected drainage systems (Figure 1). These systems were connected for measurement purposes based on former connections during glacial retreat. RRN sites and sites in the UM were connected through the headwaters of the Otter Tail and Mississippi Rivers because the RRN may have captured the Otter Tail river from the Pomme de Terre River, a tributary of the Minnesota River and the UM (Graf 1997). The Rainy River population was also connected to the UM River via Lake Winnipeg and the RRN (Graf 1997, and references therein). The Rainy River site was connected to the Lake St. Clair site via Lakes Superior and Huron. Lake St. Clair was connected to the Ohio River drainage sites (Green and Licking) via Lake Erie, the Maumee, St. Joseph, and Wabash rivers (around Ft. Wayne IN) which drained glacial Lake Erie at the end of the Pleistocene (Graf 2002). All other sites were connected via the Mississippi River.

Results

A 633-bp section of the COI gene was sequenced for 116 individuals, producing 26 unique haplotypes. Haplotypes were at most 15 bp different from each other, and there were several haplotypes that were only 1 bp different. Overall, there were no significant relationships between latitude and any measure of diversity (data not shown), Only Lake St. Clair and the Ohio River tributaries had multiple copies of the same haplotype in their populations (Table 1).

The MP gene tree was shallow and had little resolution, but *L. cardium* was monophyletic and distinct from *L. ovata* with strong bootstrap support (Figure 2). Only

two other branches had over 50% support, but one with 88% support could be considered a second lineage. This lineage had limited distribution in the Licking River, Lake St. Clair, and the MR. This lineage was also differentiated in the parsimony network, and was separated from all other sequences by three bp (Figure 3). It might be noted that haplotypes from the Green River were separated into two clades with different associations (Figure 3). In clades 1-2 and 2-2, Green River haplotypes were united with most other sites except for the Licking River. In clade 3-3, it was united with all other sites, including the Licking River.

AMOVA revealed structure among drainages, among populations within drainages, and within populations, and $F_{ST} = 0.13$. Most genetic variation (87%) was within populations ($p = 0.003$), and 4% was among populations within drainages ($p = 0.022$). Over the entire scale of this study, 9% of the variation was among drainages ($p = 0.014$). Most of the among-population variation within drainages was driven by differences between the Licking and Green Rivers of the Ohio system, and differences within the MR and its tributaries, the Boone and Minnesota Rivers (Table 2). At the largest spatial scale, the southern sites in the Licking, Ouachita, and Spring Rivers had significant pairwise F_{ST} values with nearly all of the northern sites (Table 2). The Green River population only differed from one of the UM sites and none of the other northern populations (Table 2). Despite some of the closest populations having significant pairwise F_{ST} values, there was a significant relationship between geographic and genetic distance, ($r = 0.36$, $p = 0.003$), indicating IBD and a stepping stone pattern of dispersal (data not shown). Mismatch distributions for southern populations had 2 to 4 modes (Table 1) and high variances (except for Green River) which means that southern populations were likely stable over long time periods, and there was no recent population growth or range expansion. However, in all northern populations but the Boone River, there are 2 modes with much lower variances (Table 1). Mismatch distributions of northern populations grouped together did not fit the expected distribution for a model of range expansion (Rogers & Harpending 1992), but southern populations grouped together did (Table 1). Furthermore, some of the individual northern populations fit a model of range expansion, but so did some of the southern populations (Table 1).

Closer study of the three northern systems to evaluate post-Pleistocene dispersal did not support the hypothesis of stream head capture of the Otter Tail River by the

RRN, but it did not rule out broad distribution via Lake Agassiz. Only the Pomme de Terre River population contained unique haplotypes within this smaller set of data, and the most common haplotypes were in all three systems (Figure 4). The right branch of the network is made up solely of haplotypes from the Pomme de Terre and Red rivers (Figure 4). None of the pairwise F_{ST} values was significant (Table 2). However, the magnitude of pairwise F_{ST} values between populations in the Assiniboine and Souris rivers and Pomme de Terre River are smaller than values for the Roseau and Rainy rivers with the Pomme de Terre River. The Pomme de Terre River is differentiated equally from the Roseau River and Rainy River.

Nested clade analysis is most powerful when populations are sampled intensely over the entire species range, but that was not feasible for this study. Even so, NCA was able to infer past evolutionary events for this species. Clades 2-5 and 3-2 (Figure 3) had significant outcomes which inferred past fragmentation followed by long-distance colonization or range expansion (Table 3).

Discussion

Pleistocene glaciations influenced species distributions and left evidence in patterns of genetic variation (Hewitt 1996). There was no difference in diversity between populations in the north compared to the south in *L. cardium*, but there was evidence that the Lake St. Clair population was colonized from at least two glacial refugia. The other northern populations appeared to be colonized from a Mississippian refugium. Genetic data for *L. cardium* do not support a post-Pleistocene dispersal route via the Otter Tail River, but rather widespread dispersal into most northern drainages via Lake Agassiz.

Lampsilis cardium in these populations did not exhibit a common pattern of reduced genetic diversity in northern populations as compared to southern populations. This could be due to small sample sizes for the AR populations from the Ouachita and Spring Rivers. There were only 6 samples from each of these populations; therefore, genetic diversity may be underestimated for southern populations.

Mussels in large rivers show less genetic population structure than those that live in smaller rivers or headwater streams (Berg *et al.* 1998; Elderkin *et al.* 2008; Turner *et al.* 2000), likely due to the vagility of host fish in large rivers (Berg *et al.* 2007). The pattern observed for *L. cardium* is no different; F_{ST} for these populations was 0.13,

which was within the range reported (0.031 to 0.62) for other mussels (Elderkin *et al.* 2007, 2008; Machordom *et al.* 2003). This level of population divergence can be expected if dispersal follows a stepping-stone model of gene flow, which is indicated by significant correlation between geographic and genetic distances (Slatkin 1985). Large-river species commonly found with *L. cardium*, such as *Amblema plicata* and *Actinonaias ligamentina,* also follow this model of dispersal (Elderkin *et al.* 2007, 2008). Mussels rely on fish for dispersal, and these three species use from 8-14 different hosts, with 4 common species among them in the sunfish and perch families (Mussel Host Database, The Ohio State University Museum of Biological Diversity, http://www.128.146.250.63/bivalves). These fish species are common and widespread, which would facilitate dispersal over great distances. Although smallmouth bass may move up to 75 km (Lyons & Kanehl 2002), most likely mussels are not dispersed via long-distance events, but more often over shorter distances between neighboring populations (Elderkin *et al* 2008). The timing of glochidia release may also influence dispersal distances, and both *A. ligamentina* and *L. cardium* are long-term brooders that hold glochidia overwinter, releasing them in the spring. Fish hosts common between these two mussels spawn in the spring, so reproductive migration would be occurring when glochidia are on the fish. Weak genetic structure over large spatial scales can also be recovered in species with high effective population sizes (Waples 1998); *L. cardium* is one of the more common and stable species in North America (Cummings & Mayer 1992). Large population sizes reduce genetic drift and maintain high levels of diversity, which could explain the lack of structure among these populations.

Populations of northern species are consistently shown to be the product of dispersal north from multiple glacial refugia. This is a common finding in fish species as well (Bernatchez & Wilson 1998). The strongest signal of multiple refugia for *L. cardium* is in the single Great Lakes population from Lake St. Clair. This population contained haplotypes from both lineages designated in the network and MP-tree. The smaller clade contained haplotypes from the Licking River, a tributary of the Ohio River. Other mussels in the Lake Erie drainage commonly have haplotypes that spread from an Ohio River refugium, which was connected to Lake Erie at the end of the Pleistocene via the Maumee and Wabash Rivers (Elderkin *et al.* 2008; Zanatta & Murphy 2007; Hoisington *et al*., unpublished data). The second lineage in *L. cardium* from the Great Lakes was widely spread within the Mississippi drainage. This lineage could have been isolated in

the Ozark or Ouachita highlands as it was found in both the Spring and Ouachita River populations. Nested clade analysis for the clade containing haplotypes in lineage 1 also inferred fragmentation in both highland areas, followed by dispersal into the upper Mississippi and as far north as the Pomme de Terre in MN. Dispersal from an Ozark refuge into the upper Mississippi River has also been documented for fish species, including percids (Near *et al.* 2001) which serve as hosts for *L. cardium*.

Patterns in the distribution of haplotypes in the Licking and Green Rivers could reflect the different geologic past of the upper and lower Ohio River. The parsimony network revealed some interesting groupings of haplotypes from the Ohio system. Clades 1-2 and 2-2 grouped some Green River haplotypes with most other populations, except for the Licking River. The Green River is known for its high levels of endemism (Burr & Page 1991), which would also explain the higher nucleotide diversity in this population. The rest of the Green River haplotypes grouped with some from the Licking River population and most other populations into clades 3-1 and 3-3. Prior to the Pleistocene, the Ohio River did not exist as a single river (Burr & Page 1991). The Licking River was part of the upper Ohio or Ancient Teays river system and the Green River was part of the lower Ohio that drained into the old Tennessee River. Due to the limited sampling in this study, it is not clear if there would be evidence of this deeper divergence within *L. cardium*. Green and Licking Rivers are south of the glacial maximum, but they would have been close to the edge of the ice, with fringe habitats affected by melt water. These marginal habitats in the Green and Licking Rivers probably would not have supported long-term stable populations throughout repetitive cycles of glaciers. Founder-flush cycles created as populations in these rivers were extirpated and recolonized could have caused pre-Pleistocene diversity to be lost (Bernatchez & Wilson 1998).

Results of this study do not clarify how dispersal of fauna occurred from the UM into the RRN (Assiniboine, Souris, and Roseau Rivers), LOW (Rainy River) and Lake Superior drainages. Nested clade analysis indicates there was colonization and range expansion into northern populations, but this was not demonstrated in a mismatch distribution for sequences from northern populations. Regardless, the shallow nature of the MP tree and the widespread distribution of most haplotypes are indicative of range expansion (Avise 2000). Furthermore, haplotypes from all of the above drainages are in areas formerly connected via glacial Lake Agassiz, which could have been provided the

opportunity for dispersal among all of these systems (Gagnon & Angers 2006; Rempel & Smith 1998). There were no patterns of unique haplotypes that might indicate that dispersal was only out of the Pomme de Terre River into the RRN.

Further study is needed to test dispersal of aquatic fauna from the MR into the RRN; additional sample sites would go a long way to advancing our understanding of dispersal among the RRN, LOW, and MR systems. Comparative studies of additional species would also help clarify the situation, because *L. cardium* is common among all drainages and continued gene flow in current drainage systems could complicate interpretation of genetic and geographic patterns. Studies including species found only in the RRN and UM and a different species found only in the LOW and UM drainages would provide comparative phylogenies to corroborate or refute these results.

References

Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, Massachusetts.

Berendzen PB, Simons AM, Wood RM (2003) Phylogeography of the northern hogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. *Journal of Biogeography,* **30**, 1139-1152.

Berg DJ, Cantonwine EG, Hoeh WR, Guttman SI (1998) Genetic structure of *Quadrula quadrula* (Bivalvia: Unionidae): little variation across large distances. *Journal of Shellfish Research,* **17**, 1365-1373.

Berg DJ, Christian AD, Guttman SI (2007) Population genetic structure of three freshwater mussel (Unionidae) species within a small stream system: significant variation at local spatial scales. *Freshwater Biology,* **52**, 1427-1439.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology,* **7**, 431-452.

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

Burr BM, Page LM (1991) Zoogeography of fishes of the lower Ohio-upper Mississippi basin. In: *Zoogeography of North American freshwater fishes* (eds. Hocutt CH, Wiley EO) pp. 287 John Wiley and Sons, Inc., Hoboken, NJ.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology,* **9**, 1657-1659.

Cummings KS, Mayer CA (1992) *Field guide to freshwater mussels of the Midwest*, Manual 5. Illinois Natural History Survey.

Elderkin CL, Christian AD, Vaughn CC, Metcalfe-Smith JL, Berg DJ (2007) Population genetics of the freshwater mussel, *Amblema plicata* (Say 1817) (Bivalvia: Unionidae): evidence of high dispersal and post-glacial colonization. *Conservation Genetics,* **8**, 355-372.

Elderkin CL, Christian AD, Metcalfe-Smith JL, Berg DJ (2008) Population genetics and phylogeography of freshwater mussels in North America, *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia: Unionidae). *Molecular Ecology,* **17**, 2149-2163.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes -application to human mitochondrial-DNA restriction data. *Genetics,* **131**, 479-491.

Fetzner JW, Crandall KA (2003) Linear habitats and the nested clade analysis: an empirical evaluation of geographic versus river distances using an Ozark crayfish (Decapoda: Cambaridae). *Evolution,* **57**, 2101-2118.

Gagnon MC, Angers B (2006) The determinant role of temporary proglacial drainages on the genetic structure of fishes. *Molecular Ecology,* **15**, 1051-1065.

Graf DL (1997) Northern redistribution of freshwater pearly mussels (Bivalvia: Unionoidea) during Wisconsin deglaciation in the southern Glacial Lake Agassiz region: a review. *American Midland Naturalist,* **138**, 37-47.

Graf DL (2002) Historical biogeography and late glacial origin of the freshwater pearly mussel (Bivalvia: Unionidae) faunas of Lake Erie, North America. *Occasional Papers on Mollusks Museum of Comparative Zoology Harvard University,* **6**, 175-211.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.

Hardy ME, Grady JM, Routman EJ (2002) Intraspecific phylogeography of the slender madtom: the complex evolutionary history of the Central Highlands of the United States. *Molecular Ecology,* **11**, 2393-2403.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnaean Society,* **58**, 247-276.

Lyons J, Kanehl P, eds (2002) *Seasonal movements of smallmouth bass in streams.* American Fisheries Society, Bethesda, MD.

Machordom AR, Araujo R, Erpenbeck D, Ramos MA (2003) Phylogeography and conservation genetics of endangered European Margaritiferidae (Bivalvia: Unionoidea). *Biological Journal of the Linnaean Society*, **78**, 235-252.

Mandrak NE, Crossman EJ (1992) Postglacial dispersal of freshwater fishes into Ontario. *Canadian Journal of Zoology,* **70**, 2247-2259.

Near TJ, Page LM, Mayden RL (2001) Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. *Molecular Ecology,* **10**, 2235-2240.

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

Parmalee PW, Bogan AE (1998) *The freshwater mussels of Tennessee,* first ed. The University of Tennessee Press, Knoxville, Tennessee.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology,* **9**, 487-488.

Radke S (1992) Otter Tail River fishes and mussels; an old connection. MS thesis, University of Minnesota.

Ray JM, Wood RM, Simons AM (2006) Phylogeography and post-glacial colonization patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). *Journal of Biogeography,* **33**, 1550-1558.

Rempel LL, Smith DG (1998) Postglacial fish dispersal from the Mississippi refuge to the Mackenzie River basin. *Canadian Journal of Fisheries and Aquatic Sciences,* **55**, 893-899.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution,* **9**, 552-569.

Schneider S, Roessli D, Excoffier L (2000) Arelquin, v. 2.0: a software for population genetics analysis. [Online] http://lgb.unige.ch/arlequin/software/.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial-DNA sequences in stable and exponentially growing populations. *Genetics,* **129**, 555-562.

Strange RM, Burr BM (1997) Intraspecific phylogeography of North American highland fishes: a test of the Pleistocene vicariance hypothesis. *Evolution,* **51**, 885-897.

Switzer JF (2004) Molecular systematics and phylogeography of the *Etheostoma variatum* species group *(*Actinopterygii: Percidae*).* PhD thesis, Saint Louis University.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Molecular Biology and Evolution,* **10**, 512-526.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution,* **24**, 1596-1599.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history — a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics,* **140**, 767-782.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology,* **7**, 381-397.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology,* **13**, 789-809.

Turner TF, Trexler JC, Harris JL, Haynes JL (2000) Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. *Genetics,* **154**, 777-785.

Vucetich JA, Waite TA (2003) Spatial patterns of demography and genetic processes across the species' range: null hypotheses for landscape conservation genetics. *Conservation Genetics,* **4**, 639-645.

Walker JM, Curole JP, Wade DE, Chapman EG, Bogan AE, Watters GT, Hoeh WR (2006) Taxonomic distribution and phylogenetic utility of gender-associated mitochondrial genomes in the Unionoida (Bivalvia). *Malacalogia,* **48**, 265-282.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity,* **89**, 438-450.

Wares JP, Turner TF (2003) Phylogeography and diversification in aquatic mollusks. In: *Molecular systematics and phylogeography of mollusks* (ed. Lydeard, C. Lindberg, D.R.) pp. 229-269 Smithsonian Institution, Washington D.C.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population structure. *Evolution,* **38**, 1358-1370.

Whiteley AR, Spruell P, Allendorf FW (2004) Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. *Molecular Ecology,* **13**, 3675-3688.

Zanatta DT, Murphy RW (2007) Range-wide population genetic analysis of the endangered northern riffleshell mussel, *Epioblasma torulosa rangiana* (Bivalvia: Unionoida). *Conservation Genetics,* **8**, 1393-1404.

Zanatta DT, Murphy RW (2008) The phylogeographical and management implications of genetic population structure in the imperiled snuffbox mussel, *Epioblasma triquetra* (Bivalvia: Unionidae). *Biological Journal of the Linnaean Society,* **93**, 371-384.

Table 1. Descriptive statistics of sampling sites for *Lampsilis cardium.*

Site Name	Site ID	N	Corrected number of haplotypes	Number of unique haplotypes	Mean no. base pairs different	Nucleotide diversity (π)	Mismatch modes	Mismatch variance
Northern	-	-	-	-	-	-	3	66.7
Upper Miss. Pool 8	M1	22	22	1	3.10	0.0049	3*	4.1
Upper Miss. Pool 14	M2	4	4	0	4.67	0.0074	2	13.1
Pomme de Terre River	PT	14	14	1	4.72	0.0084	3	285.9
Assiniboine River	AS	4	4	0	2.00	0.0032	2	2.4
Souris River	SR	5	5	0	2.60	0.0041	2	2.0
Roseau River	RO	4	4	0	1.50	0.0024	2*	2.7
Rainy River	RR	6	6	0	1.93	0.0031	2*	2.4
Boone River	BR	5	5	1	1.80	0.0028	1*	1.5
Lake St. Clair	LS	8	5	2	4.64	0.0073	4*	5.2
Southern	-	-	-	-	-	-	4*	122.6
Spring River	SP	6	6	1	2.00	0.0032	2*	1.7
Ouachita River	OR	6	6	0	1.20	0.0019	2	1.6
Licking River	LR	14	14	2	7.86	0.0124	4*	116.7
Green River	GR	18	14	6	7.67	0.0120	4*	201.4

*Indicates mismatch distributions fit a model of range expansion.

Table 2. Pairwise F_{ST} estimates (Tamura & Nei 1994) for the partial COI gene in 13 populations of *Lampsilis cardium*. Locations are as in Table 1.

Location	M1	M2	PT	AS	SR	RO	RR	BR	LS	SP	OU	LR
M1												
M2	0.2347*											
PT	0.0442*	0.0626										
AS	0.0879	0.2593*	0.0349									
SR	0.0431	0.2102	0.0066	-0.2166								
RO	0.2351*	0.3151	0.0858	-0.1667	-0.0657							
RR	0.0772	0.3544*	0.0784	-0.1914	-0.1174	0.1112						
BR	0.1211	0.3084*	0.0730	-0.2644	-0.1457	-0.0693	-0.1574					
LS	0.0398	-0.0949	0.0273	0.0308	-0.0011	0.1247	0.0839	0.0679				
SP	0.1476*	0.4337*	0.0144	0.4801*	0.3798*	0.595*	0.4789*	0.4996*	0.212*			
OU	0.3230*	0.5997*	0.0998	0.6633*	0.5824*	0.7501*	0.6510*	0.6662*	0.3849*	0.2613*		
LR	0.1929*	0.2024*	0.1157*	0.2169*	0.201*	0.2886*	0.2359*	0.2425*	0.1607*	-0.0087	0.1232	
GR	0.0538*	0.1429	0.0117	0.0478	0.0335	0.1144	0.0778	0.0872	0.0627	-0.0494	0.0903	0.0423*

Estimates are significantly greater than zero (*) at $p = 0.05$ using 1023 nonparametric permutations of the data.

Table 3. Nested clade analysis of geographical associations for clades from Figure 3.

Clade	Permutational chi-square statistic	Probability	Inference Key # and outcome
2-5	40.53	0.0004	1, 2, 3, 5, 15, past fragmentation and long distance colonization
3-2	137.01	0.0000	1, 2, 11, 12, 13, fragmentation and then range expansion

Clades with neither genetic nor geographic variation and clades without enough geographic sampling to make inferences were excluded.

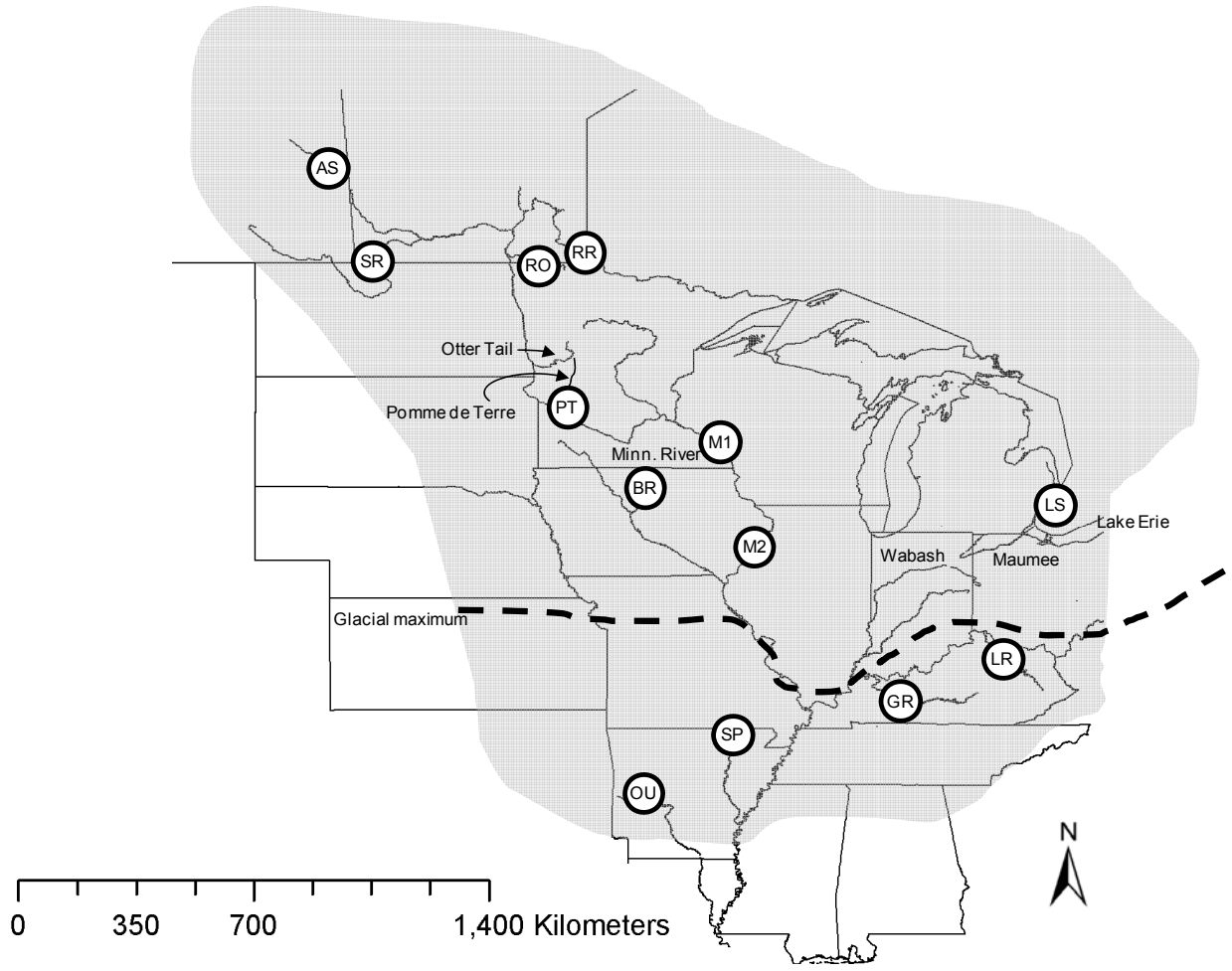

Figure 1. Distribution range for *Lampsilis cardium* in North America (shaded region) and sampling locations (circles) for phylogeographic analyses. Location labels as in Table 1.

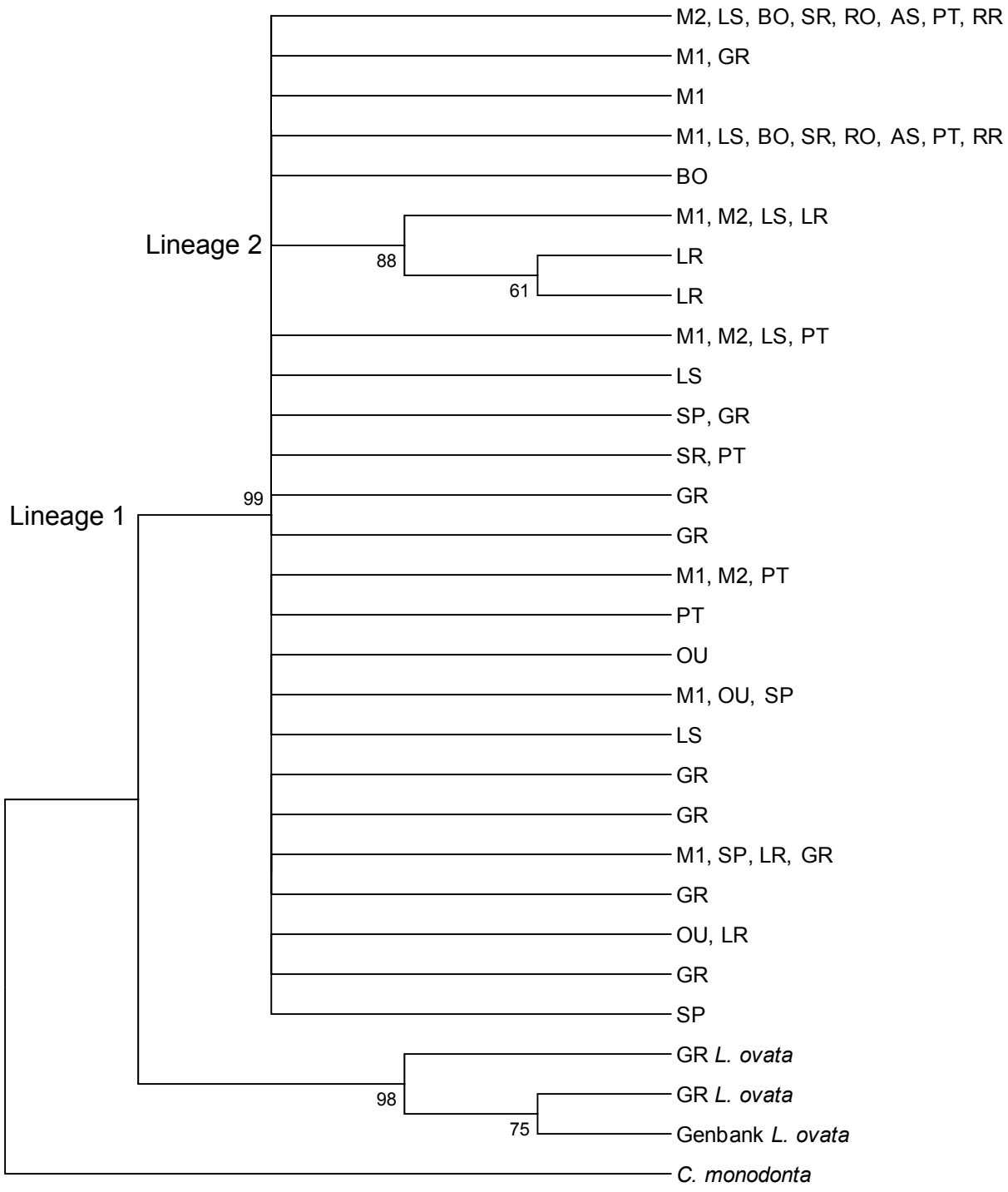

Figure 2. Maximum parsimony tree for *Lampsilis cardium* partial COI sequences. Locations where haplotypes were found follow Table 1. Numbers at nodes indicate percent support from 1000 bootstrap replications. Outgroups discussed in text.

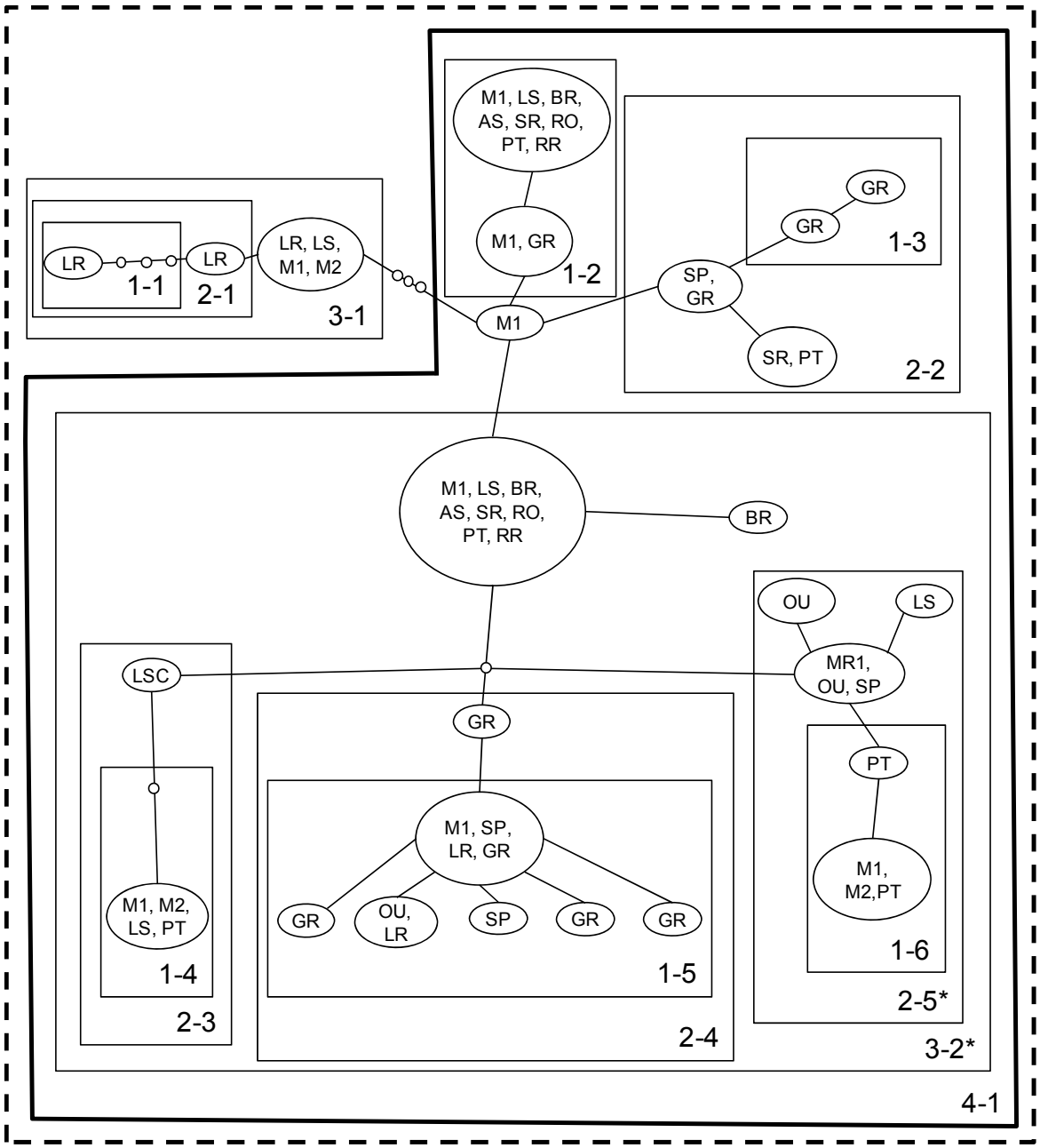

Figure 3. Parsimony network with nested clades (boxes) for *Lampsilis cardium* COI haplotypes showing 1-, 2-, 3-, and 4-step clades that contain genetic and geographical variation. Lines connecting ovals represent one base-pair difference between haplotypes (ovals), small circles are inferred missing haplotypes, and haplotype frequency is relative to the size of the oval. Lineage 2 is contained within clade 3-1. Asterisk (*) indicates significant inference in NCA. Letters designate locations where the haplotype was found, abbreviations follow Table 1.

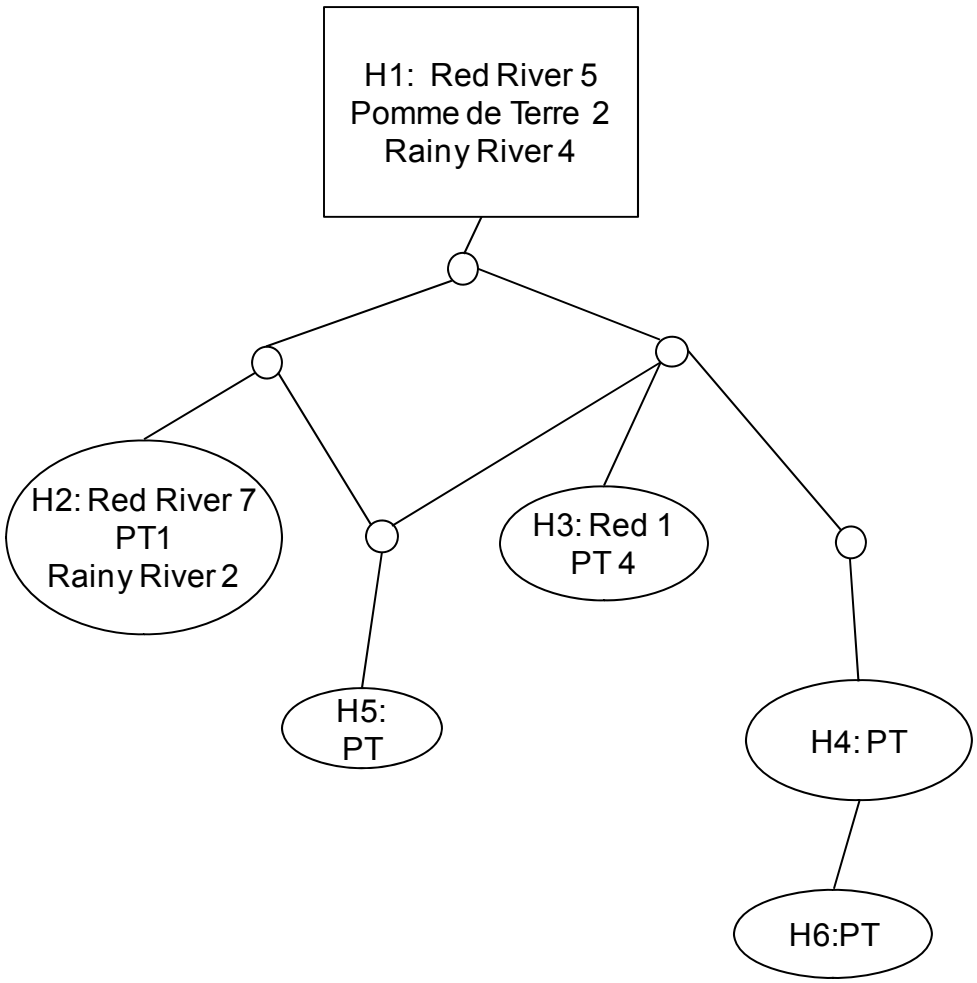

Figure 4. Parsimony network for *Lampsilis cardium* COI haplotypes from Pomme de Terre River, Red River tributaries and Rainy River populations. Lines represent one base-pair difference between haplotypes represented as ovals, small circles are inferred missing haplotypes, and haplotype frequency is relative to the size of the ellipse. Square haplotype is the 95% confidence ancestral haplotype. PT is Pomme de Terre.

Chapter 4

Population genetic analysis of *Lampsilis cardium* in the impounded upper Mississippi River

Abstract

Impoundment of freshwater habitats creates barriers to dispersal for mussels by blocking movement of host fish, and therefore may reduce gene flow among mussel populations. Highly variable microsatellite markers are effective in revealing fine-scale population structure in marine and freshwater organisms. Eight microsatellite markers were used to assess genetic variation in *Lampsilis cardium* within and among populations in river reaches of the upper Mississippi River (UMR), which is fragmented by a system of locks and dams. Assignment tests were used to estimate contemporary levels of gene flow for comparison to historic levels inferred with F-statistics. If dams are effective barriers to mussel dispersal and gene flow, then contemporary gene flow should be lower than historica gene flow. Most genetic variation was contained at the population level, and there was no structure evident within the UMR. F_{ST} values were non-significant among populations in the river, and even at broader spatial scales encompassing more distant populations; only weak structure was found. This stretch of the UMR may be fragmented for this species. Contemporary estimates of gene flow appear to be lower than estimates of historic gene flow. Tests for equilibrium between gene flow and genetic drift indicate that high gene flow occurred historically for this species explaining the lack of genetic structure found in *L. cardium*.

Introduction

Freshwater river systems are some of the most anthropogenically altered ecosystems (Strayer 2006) having been physically modified via channelization and impoundment. In North America, 72% of all large river systems are moderately or heavily affected by fragmentation, channelization, or impoundment (Nilsson *et al.* 2005). Altering the physical nature of rivers destroys habitats, such as fast-flowing water, and cuts off main-channel areas from side channels or back water lakes. Reservoirs created behind dams submerge shallow river habitats in deep, slow moving water, displacing benthic organisms and changing pelagic species due to changes in light penetration, nutrient dynamics, and food web arrangements (Amoros & Bornette 2002). Dams also fragment river systems into shorter river reaches, which negatively affects some aquatic fauna (Nilsson *et al.* 2005; Pringle 2001).

Habitat fragmentation can be defined as the breaking apart of habitat (Fahrig 2003), which isolates patches of suitable habitat within a matrix of unsuitable habitat (Wiegand *et al.* 1995). In lotic systems, upstream habitats are separated from downstream habitats by dams (Stanford & Ward 2001). Dams are also barriers to dispersal for vagile aquatic organisms, and even low-head dams of less than one meter can prohibit fish movement upstream (Watters 1996). Fragmentation of rivers affects distribution, growth, and abundance of fish (Quinn & Kwak 2003). When dispersal is prevented, gene flow and connectivity among populations are stopped. Dams also reduce genetic diversity in some fish species (Haponski *et al.* 2007; Hanfling & Weetman 2006; Yamamoto *et al.* 2004), but not others (Reid *et al.* 2008). Fish and unionid mussels are tightly linked in nature, because fish act as hosts for parasitic larvae (glochidia) of mussels. Mussel dispersal over longer distances, especially upstream, primarily occurs as larvae travel on host fish. Adult mussels are relatively sessile, while gametes and juveniles are primarily dispersed downstream with flow (Parmalee & Bogan 1998).

Freshwater mussels have experienced precipitous declines due to impoundment of rivers (Lydeard *et al.* 2004). Mussel species that only use one or a few host fish have been especially affected by dams impeding dispersal of their hosts (Watters 1996). For example, the elephant ear (*Elliptio crassidens*, Lamarck) and ebony shell (*Fusconaia*

ebena, Lea) mussels use skipjack herring (*Alosa chrysochloris,* Rafinesque) as hosts, and populations of these mussel species above dams that block fish passage upstream (Becker 1983) are slowly becoming extirpated as old animals die and no recruitment occurs (Kelner 2003). Impoundments also create extinction gradients for mussels in river reaches below dams (Vaughn & Taylor 1999). Low-head dams affect the distribution of mussel species in the midwest (Watters 1996) and south (McGregor & Garner 2004). Not all species of mussels show reduced dispersal and abundance however. A recent study of *Lampsilis cariosa* did not find any relationship between genetic diversity and the number, age, or height of dams (Kelly & Rhymer 2005), perhaps because the study covered a large area in three drainage systems with limited sampling in each system (Kelly & Rhymer 2005).

Understanding population genetic structure and gene flow for mussel species is imperative to conserving this highly endangered fauna (National Native Mussel Conservation Committee1998). This is true even for species that are relatively healthy because they are in decline from a variety of threats (Berg *et al*. 2008; Lydeard *et al.* 2004). There is a small, but growing body of information on genetic variation in freshwater mussels. However, much of this work has focused on resolving phylogenetic and phylogeographic relationships (King *et al.* 1999; Mulvey *et al.* 1997; Roe *et al.* 2001; Roe & Lydeard 1998a; Roe & Lydeard 1998b; Serb *et al.* 2003; Zanatta & Murphy 2006), or on describing genetic structure across broad geographic ranges (Berg *et al.* 1998; Kelly & Rhymer 2005; Machordom *et al.* 2003). Relatively few studies have examined small-scale, local genetic variation within a species of mussel (but see Berg *et al.* 2007). Additionally, most of these studies have used mitochondrial (mt) gene sequence markers, which are most useful at larger spatial scales or for inferring evolutionary history (but see Geist & Kuehn 2005; Kelly & Rhymer 2005). Most allozyme and mtDNA studies have not revealed small-scale or local genetic variation for several aquatic organisms (Duran *et al.* 2004b; Wirth & Bernatchez 2001), including mussels (Berg *et al.* 1998, but see Berg *et al.* 2007), whereas microsatellite markers have revealed small-scale, local variation within and among populations in some of these same organisms in later studies (Duran *et al.* 2004a; Shaw *et al.* 1999), including mussels (Geist & Kuehn 2005). Because of their high variability, microsatellites are the

single most informative method of current molecular techniques for measuring DNA variation within and among populations (Avise 2000, 2004; Frankham 2003; Page & Holmes 1998), and the most useful markers to infer recent gene flow among populations of organisms.

To study the effects of habitat fragmentation on gene flow, information is needed on past and current levels of gene flow. Dispersal of individuals among populations can be directly monitored in the field, but it is difficult to confirm reproduction, and hence gene flow in the field. Information about population ecology can be gathered with genetic analyses that are qualitatively similar to that obtained with traditional field techniques (Paetkau *et al.* 2004). Microsatellite data is variable enough to provide information about individual organisms (Estoup & Angers 1998), and simulation studies have shown that it is possible determine direct, real-time estimates of dispersal (Paetkau *et al.* 2004) using assignment tests. Agreement between assignment tests and genetic data has also been confirmed with field observations (Berry *et al.* 2004). Furthermore, since assignment tests use genetic data these dispersal estimates also include measurement of gene flow, a parameter that cannot be measured by simply monitoring movement of individuals among populations with field observations. Ecological study of mussels in large rivers is difficult due to deep, fast water. The long lifespan of mussels (some live 100 years) makes it impractical to conduct meaningful mark and recapture studies within a researcher's lifetime to collect dispersal information. Mark and recapture studies of mussels are further hindered by vertical dispersal of mussels within the substrate (Amyot & Downing 1997; Watters *et al.* 2001), which can lead to reduced recapture rates for several species. Thus, mussels are good candidates for study with genetic data.

Lampsilis cardium is a widespread mussel species that lives in medium and large rivers, and is considered stable across much of its range (Natureserve 2008). Individuals live at least 50-100 years, become sexually mature around 7-8 years (T. Newton, personal communication) and use several species of host fish including; largemouth bass (*Micropterus salmoides*, Lacépède), smallmouth bass (*M. dolomieu*, Lacépède, bluegill (*Lepomis macrochirus*, Rafinesque), sauger (*Sander canadensis,*Griffith & Smith), walleye (*S. vitreus,* Mitchill), white crappie (*Pomoxis*

annularis, Rafinesque) and yellow perch (*Perca flavescens*, Mitchill) (Watters & O'Dee 1999; Mussel Host Database, The Ohio State University Museum of Biological Diversity, http://www.128.146.250.63/bivalves). Some of these hosts move greater distances upstream and downstream (e.g. walleye) than others (e.g. bluegill; Maracek & Wlosinski 1996).

The highly fragmented Upper Mississippi River (UMR) is that portion of the Mississippi River controlled by a system of locks and dams stretching from St. Paul, MN, to Cairo, IL. The dams separate distinct river reaches called navigation pools, which creates a framework for measuring the population genetic variation of mussels within a spatial hierarchy. The goal of this study was to evaluate the effect of fragmentation on mussel populations along a 430-km stretch of the UMR. One objective was to determine if there was population subdivision or population differentiation in navigation pools separated by dams. The second objective was to determine if there are lower levels of contemporary gene flow compared to historic gene flow. To meet objective one, genetic diversity was measured within a spatial hierarchy — within populations, among populations within pools, and among pools (Figure 1). To meet objective two, historic gene flow was inferred and compared to contemporary gene flow. Assignment tests were used to estimate contemporary gene flow (Berry *et al.* 2004; Paetkau *et al.* 2004), while historic gene flow was estimated traditionally as the number of migrants (Nm) derived from F_{ST} (Whitlock & McCauley 1999). If habitat fragmentation has created barriers to gene flow among populations or beds of mussels, then genetic diversity should be partitioned within and among pools and there should be differences in historic versus contemporary gene flow.

Materials and methods

Sample collection

Lampsilis cardium were obtained from populations in 6 pools of the UMR. In the UMR, mussels were sampled from a total of 16 populations — three pools (4, 11, 14) were represented by a single population each and three pools (7, 8, 9) were represented by multiple populations (Figure 1). For comparison at much broader spatial scales, populations greater distances away were also sampled: one from the Boone

River, IA and one from Lake St. Clair, part of the Great Lakes drainage system. Approximately 20 mussels from each population were collected by hand while wading or diving, a small clip of mantle tissue was removed non-lethally (Berg *et al.* 1995), and mussels were returned alive to the river bottom. Mussels from the UMR were measured for length to the nearest mm. Tissue samples were preserved on dry ice or in 95% ethanol in the field and stored at -80°C upon return to the lab.

Molecular analyses

Total genomic DNA was extracted from mantle tissue using DNeasy kits (Qiagen, Valencia, CA) or by phenol/chloroform (Sambrook *et al.* 1989). Thirteen of 15 published primer pairs for a congener (Eackles & King 2002) were screened, and PCR components and protocols were optimized for target allele amplification. Allele sizes within the ranges reported (Eackles & King 2002) were considered target alleles. Eight primers (LabD111, LabD10, LabD206, LabC67, LabD213, LabD29, LabC24, and LabC2; hereafter referred to only by their numbers) consistently amplified alleles in *L. cardium*. PCR amplifications consisted of 100-200 ng of genomic DNA, 1x PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 2.5 mM (24 and 2), 3.0 mM (primers 111, 10 and 29), or 3.5 mM (primers 206, 67, 213) $MgCl_2$, 0.15% Tween-20, 0.25 mM dNTPs, approximately 0.5 µM forward and reverse primer, and 0.4 U *Taq* DNA polymerase in a total reaction volume of 10 or 15 µL.

Amplifications were carried out on PTC-200 or 100 DNA Engine Cyclers (MJ Research, Waltham, MA) using the following conditions: 94°C initial heating for 2 minutes followed by 45 cycles at 94°C for 40 seconds, 40 seconds at annealing temperatures (50°C for 2 and 29; 53°C for 24; 56°C for 111, 10, and 206; and 60°C for 67, 219, and 29), and extension at 72°C for 1 minute, concluding with a final extension at 72°C for 30 minutes to eliminate split-peak artifacts and to push all fragments to the plus-A length (Applied Biosystems 2005). Visualization of PCR products was conducted on an Applied Biosystems (ABI; Foster City, CA) 3730 automatic genetic analyzer that used capillary electrophoresis to resolve DNA fragments with fragment size calls made in GeneMapper software (ABI). Subtle differences in capillary runs and point mutations in DNA can create allele fragments that are not exact multiples of microsatellite base pair repeats which may artificially inflate diversity estimates (Amos *et*

al. 2007). Therefore, allele fragment sizes were sorted into allele categories using an Excel add-in macro, Flexibin (Amos *et al.* 2007). This method uses an algorithm for binning alleles that maximizes allele calling accuracy and prevents overestimating the number of alleles. Microchecker (van Oosterhout *et al.* 2006) was used to check binned alleles for genotyping errors, especially the presence of null alleles which are common in mussel genomes (Kelly & Rhymer 2005).

Data analyses

Summary statistics were calculated with GenAIEx software (Peakall & Smouse 2006). Allelic richness was calculated after adjusting for uneven sample sizes using rarefaction (Petit *et al.* 1998) as implemented in Fstat (Goudet 1995). Genepop (Raymond & Rousset 1995b) was used to test for deviation of genotype frequencies from Hardy-Weinberg expectation and to conduct randomization tests for linkage disequilibrium. Population differentiation and structure were assessed with several different methods. First, Weir and Cockerham's (1984) F_{ST} -analog (θ) was computed for all pairs of populations within the UMR and for pairs of pools and the Boone River and Lake St. Clair populations in Arlequin (Schneider *et al.* 2000). Next, exact tests for population differentiation were conducted in Genepop (genic differentiation option) for the UMR populations at two different levels, population and pool. Genic differentiation tests the null hypothesis that allelic distribution is identical across groups; this was run with using the Markov chain Monte Carlo method in 100 batches with 1000 iterations per batch for both levels. Results of locus-specific tests were combined (Fisher's method) to give an overall test of differentiation among populations and among pools because this method gives more weight to rare alleles and is more sensitive to weak population differentiation than F_{ST} (Raymond & Rousset 1995a). Finally, analysis of molecular variation (AMOVA; Excoffier *et al.*1992) was used in Arlequin at two different spatial scales. Two separate runs were conducted due to the limitation of levels allowed by the AMOVA analysis in Arelquin (Figure 1). To measure variation within the UMR only, AMOVA partitioned variation into three levels including within populations, among populations within pools, and among pools. A second AMOVA was conducted to include variation at the broadest spatial scale which partitioned variation into three levels including within populations, among populations within drainages, and among

drainages. F-statistics were used to estimate gene flow where $F_{ST} = 1/(4Nm+1)$ (Slatkin 1995; Wright 1978). To separate the influence of gene flow from drift on the current population structure, I used Hutchinson and Templeton's (1999) graphical approach that examined the correlation of F_{ST} with geographic distance. Residuals of F_{ST} and geographic distance (km) were calculated with regression in Excel, and a manual Mantel test of the residuals and pairwise F_{ST} matrices was run in Arlequin.

Estimating gene flow is typically based on Wright's F-statistics, which can be biased because natural populations often do not meet the assumptions of the model (Pearse & Crandall 2004). Furthermore, F_{ST} tends to reflect past levels of gene flow (Whitlock & McCauley 1999). In recent years, there has been an explosion of new analyses for inferring dispersal. A useful application of this method is to compare historic versus contemporary gene flow (Paetkau *et al.* 2004; Pearse & Crandall 2004). Assignment tests can determine the most likely source of an individual among the sampled populations (Paetkau *et al.* 2004). The assumptions vary among different assignment tests (Pearse & Crandall 2004), and in this study, not every possible mussel population in this stretch of the UMR was sampled. Therefore, I used the assignment test in Geneclass2 (Piry *et al.* 2004), which did not require all possible populations to be sampled, to calculate the likelihood that a given multi-locus genotype originated from any of the sampled populations. The Bayesian criterion method (Rannala & Mountain 1997; Paetkau *et al.* 2004) was implemented with 1000 simulations and $\alpha = 0.05$. Assignment tests were conducted at two levels to assess the affect of habitat fragmentation on gene flow. First, analyses were conducted at the population level, which would indicate gene flow prior to habitat fragmentation, and a second assignment test was performed with populations grouped into pools which would reflect fragmentation from dams. An increase in assignment rates from the population level to the pool level would be consistent with the hypothesis that dams are reducing gene flow.

Isolation-by-distance was tested by comparing pairwise genetic distances (F_{ST}) to geographic distances with a Mantel test in Genepop. Geographic distances between each pair of populations were measured to the nearest km using the path measurement tool in GoogleEarth© (www.googleearth.com) by following each bend in the river.

Populations were placed into GoogleEarth with coordinates obtained onsite with a Garmin (Olathe, KS) handheld Global Positioning Unit. Correlation between genetic distance and the number of dams between populations was tested with a Mantel test in Genepop. Because multiple populations were sampled in some pools, to avoid values of zero in the matrix, one was added to each pairwise value for the number of dams.

The clustering method in Partition (Dawson & Belkhir 2001) was used to determine how many populations were sampled over all locations. Two separate runs were conducted with 10,000 Markov chain Monte Carlo observations collected every 10 iterations and results were evaluated after 3000 burn-in steps. Priors were set to 18 maximum populations (Boone, Lake St. Clair, and each population in the UMR), θ was set at the minimum value allowed 0.001, and μ was set to 1 (Dawson & Belkhir 2001).

Results

Genetic diversity

All eight loci amplified in *L. cardium* were polymorphic, with similar size ranges to those reported for other *Lampsilis* species (Eackles & King 2002; Kelly & Rhymer 2005). These eight loci had a total of 213 alleles and averaged 26 alleles per locus (range 6-43; Table 1). Allelic richness and levels of heterozygosity (H) were similar in each population within the UMR, and both distant populations; richness ranged from 8.7-9.4, and H ranged from 0.77 to 0.85 (Table 1). Departure from Hardy-Weinberg expectations (HWE) was significant in nearly half of all locus-by-population combinations for all 18 populations (66 of 144; Table 1). More powerful tests for heterozygote deficiencies were conducted for only the UMR populations, and 45% of 144 all locus-by-population combinations were significant. Microchecker revealed that in most cases, deviation from HWE was likely due to null alleles, because there was no indication that other genotyping errors, such as large allele drop out, occurred. Frequencies for null alleles in each locus-by-population combination were variable, only one combination had a high frequency of 0.33, and the average was low at 0.05. Ten out of 144 locus-by-population combinations tested significant for linkage disequilibrium, but this was scattered among loci in different populations. There were no significant

linkages among loci; therefore, each locus will be considered an independent measure of diversity.

Population structure

Pairwise F-statistics revealed very little genetic differentiation among populations in the UMR: only four had significant values, and three of these included population three in pool 7 (Table 2). Combining populations into pools within the river yielded one significant pairwise F_{ST} difference between neighboring pools 7 and 8 (Table 3). Exact tests revealed a little more differentiation at both the population (Table 2) and pool (Table 3) levels. There were 13 significant differences between population pairs in the UMR, and nine at the pool level. However, despite these significant differences, overall differentiation among populations was very low. Pairwise F_{ST} values were not over 0.046, and in the exact tests, only four of eight loci were different from zero at either the population (Table 2) or pool (Table 3) levels. This weak differentiation within the UMR was compared to values for spatial structure at a much broader scale using populations in the Boone River and Lake St. Clair. At this scale, pairwise F_{ST} values revealed that the population from the Lake St. Clair was significantly different from four of the UMR pools as well as the Boone River population (Table 3). However, the Boone River was only differentiated from two UMR pools (Table 3).

Weak, but significant, structuring of populations as measured by θ was found within populations, but not among populations or among pools. Within the UMR only, θ = 0.003, and only 0.3% of the total variation was partitioned among populations with a majority of variation (99%) contained within populations (Table 4a). Even at the largest spatial scale, including the Boone River and Lake St. Clair populations, θ = 0.011 which indicated there was very weak but significant (p = 0.04) structure among drainages (Table 4b).

Gene flow

Estimates of the number of migrants per generation using F_{ST} were high, and averaged 373 over all populations in the UMR. The lowest estimate was 27, but the highest estimate was 7,600 (data not shown). Tests for equilibrium between gene flow and genetic drift (Hutchison & Templeton 1999) revealed no significant relationship

between genetic and geographic distance (Figure 2a); therefore, equilibrium was CASE II (Hutchison & Templeton 1999). CASE II indicates gene flow is strong relative to genetic drift. With high levels of gene flow and lack of population structure, assignment rates were very low at the population level within the UMR, where only 10% of the mussels were correctly assigned to the population from which they were sampled (Table 5). There was an increase in assignment rate at the pool level to 30% (Table 5). There was no significant relationship between genetic (F_{ST}) and geographic distances or genetic distance and the number of dams between populations in the UMR, indicating no isolation-by-distance (Figure 2a and b).

Because there was a lack of structure and high gene flow among populations of *L. cardium* in the UMR, these data were used to determine how many populations were sampled over all locations with a clustering method in Partition. Both runs overwhelmingly supported one population (75% posterior distribution), with weaker support for 2 populations (23% posterior distribution).

Discussion

Similar to other species of mussels, genetic diversity in *L. cardium* is high, but in contrast to other mussel species, populations in the UMR may be panmictic. There is weak evidence that fragmentation may be reducing gene flow for this species, but it is likely that not enough time has passed since the dams were constructed to create strong differences in genetic structure among these populations. For a distance of 430 km along the Mississippi River, there was no population structure, no differentiation among populations, but there was differentiation among UMR populations and distant populations in the Boone River and the Lake St. Clair. Tests for equilibrium between gene flow and drift indicate that levels of gene flow have been high for this species, which would lead to the lack of genetic structure observed in this study.

Genetic diversity

Levels of genetic diversity in *L. cardium* are similar to other freshwater mussels examined for microsatellite markers. High diversity, measured by allelic richness, was comparable with *L. cariosa* from Maine (Kelly & Rhymer 2005), but was higher than *Epioblasma* species (Zanatta & Murphy 2007, 2008) and *L. fasciola* (Zanatta *et al.*

2007). The lower values reported for *Epioblasma* species and *L. fasciola* may be due to smaller sample sizes. The *Epioblasma* species are threatened due to small population size, and the study on *L. fasciola* was conducted to examine sympatric speciation, rather than population structure. Populations within the UMR and at broader spatial scales had similar levels of allelic richness, even in the Boone River which is considerably smaller than the UMR. This suggests that populations are large enough to maintain high levels of genetic diversity in both medium and large rivers of the midwest. Furthermore, dams do not appear to have adversely affected genetic diversity in L. *cardium* in the UMR. This could be because the dams may not have been in place long enough —most were constructed in the 1930's and 40's. Given their long life spans, there may have only been about 10 generations of mussels since dam placement.

Since mussels are tightly linked to fish, genetic studies on fish in highly fragmented rivers are useful for comparison. A recent study on black redhorse revealed substantial genetic diversity and lack of differences among populations above or below dams (Reid *et al.* 2008) in a similarly fragmented river. These results are in contrast with other species of fish that have lowered genetic diversity in fragmented systems. In those studies, reduced diversity has been attributed to small river reaches, which decreases population size, and therefore genetic diversity (Frankham 1996; Jager *et al.* 2001). In the UMR, navigation pools in the study region are rather long (mean length = 40 km; range 14-75km), and *L. cardium* populations remain large (Kelner 2003). Recent surveys estimate the population of *L. cardium* in pool 5 of the UMR is around three million (range one million to 3 million; T. Newton, unpublished data). Large populations are able to maintain high levels of genetic diversity (Wright 1978).

It is not unusual for mussel populations to deviate from Hardy-Weinburg equilibrium at microsatellite loci (Kelly & Rhymer 2005; Zanatta & Murphy 2007, 2008; Zanatta *et al.* 2007). Heterozygote deficiencies in *L. cardium* populations may have biological explanations, such as population admixture or inbreeding, but it is more likely due to the presence of null alleles which are common in bivalves (Kelly & Rhymer 2005; Zanatta & Murphy 2007; Zanatta *et al.* 2007). These null alleles should not affect the outcome of further analyses because the frequency of null alleles in this study was very low at 5%. Low levels of null alleles did not change results for *L. cariosa* (Kelly &

Rhymer 2005), or in simulated data sets where null frequencies were as high as 20% (Dakin & Avise 2004).

Population structure and gene flow

At the broadest spatial scale covered in this study, θ was low (0.011) but significant. This indicates weak population structure, as does the significant pairwise F_{ST} values between UMR populations and both the Boone River and Lake St. Clair populations. However, because there was only one population sampled from both the Boone River and the Lake St. Clair, the variation among populations within drainages measured is driven by UMR populations. Therefore, these results should be interpreted with caution, and additional study is recommended at broad spatial scales for this species.

Results within the UMR are similar to other studies in mussels because most of the population structure revealed by microsatellites in mussels shows structure among drainages (Kelly & Rhymer 2005; Zanatta & Murphy 2007, 2008, Zanatta *et al.* 2007), but not much structure within a river system. In the UMR, there was no population differentiation, θ was near zero and pairwise F_{ST} values for these mussel populations were very low; the highest value was 0.046. In comparison, F_{ST} among drainages of *L. cariosa* was 0.04, and the pairwise F_{ST} values ranged from 0.009 to 0.185 (Kelly & Rhymer 2005). Pairwise F_{ST} values for other species ranged from -0.0013 to 0.385 (Zanatta & Murphy 2007, 2008; Zanatta *et al.* 2008). There are very few studies on mussel populations over this fine spatial scale within a river, although mussels in the Ohio River had significant yet weak population structure over long distances (1000+ km) measured with allozyme markers (Berg *et al.* 1998). Collectively, these results for several mussel species indicate high gene flow in large rivers.

Gene flow is a likely explanation for results for *L. cardium* because tests for equilibrium between gene flow and genetic drift were CASE II (Hutchison & Templeton 1999), meaning gene flow is much more prevalent than genetic drift for populations in the UMR. A second indicator of high gene flow is the majority of genetic variation is contained within populations, which has also been found in other species of mussels within these same river systems (Zanatta & Murphy 2008). A third indicator of high gene flow is the lack of correlation of F_{ST} with either geographic distance or the number

of dams between populations, and therefore a lack of isolation-by-distance (IBD). IBD indicates dispersal and gene flow via a stepping-stone model where populations that are nearest-neighbors exchange genes more often than populations that are greater distances apart (Slatkin 1985). Freshwater mussels living in smaller streams may be the ideal organism dispersed via the stepping-stone model because they are distributed in distinct patches within linear river systems. Several species of mussels have shown highly significant IBD (Kelly & Rhymer 2005; Zanatta & Murphy 2007, 2008; Zanatta *et al.* 2007), even if the magnitude of the relationship was weak (Berg *et al.* 1998; Elderkin *et al.* 2007). However, mussel populations in the UMR may not fit the stepping-stone model because gene flow is high in this system, and high rates of dispersal may occur among populations that are not nearest-neighbors. Dispersal further than the closest populations may occur for *L. cardium* based on timing of glochidia release. Gravid *L. cardium* usually keep fertilized glochidia in their gills over winter, but some females may release glochidia in the fall, to overwinter on fish hosts. This strategy may increase the distance larvae are dispersed because they are on fish for 4-6 months (Watters & O'Dee 1999).

High gene flow has also been found in other aquatic species. For example, high gene flow was inferred for black redhorse (*Moxostoma duquesnei,*Lesueur) populations, based on low population differentiation in the fragmented Grand River of Canada (Reid *et al.* 2008). High gene flow may be possible for these fish due to traits such as group spawning and long migration distances between winter habitats and spawning grounds (Reid *et al.* 2008). Similarly, traits such as reproductive timing in *L. cardium* may also promote gene flow, even with dams in place within the UMR. This species generally releases their young when water temperature is around 19°C (Holland-Bartels & Kammer 1989), which is often coincident with spring flooding; this corresponds to the time and dams are open. Although a review of 84 fish movement studies in the UMR found dams do adversely affect fish movement (Wlosinski & Marecek 1996) because only 13% of 5253 recaptured fish moved either up or downriver, there were several hundred fish that moved upriver through dams. Fish moving upriver in these studies included host species for *L. cardium* such as walleye (484), sauger (160), largemouth bass (11) and smallmouth bass (4) (Wlosinski & Marecek 1996). This study also

showed that high flow, commonly reached during open river conditions, precludes fish movement upstream. However, there were still a few fish that moved upstream during high water (Wlosinski & Marecek 1996). Therefore, occasional fish movement upriver during the time *L. cardium* release glochidia may be enough to maintain population similarity because theoretical study has shown that just one effective migrant per generation can prevent population differentiation (Wang 2004).

The results of this study support panmixia for *L. cardium* within the UMR because F_{ST} is nearly zero, and there are very high estimates of the number of migrants per generation, and Partition indicated one population was sampled in this study. This strongly suggests that low population differentiation has been the condition for this species historically, and it has evolved with high levels of gene flow over great distances. The flood-pulse concept (Junk *et al.* 1989) of river systems may best explain historical levels of high gene flow in the UMR. This concept suggests that large, flood-plain rivers such as the UMR act as super-highways for fish, promoting movement between habitats that provide resources for feeding and/or breeding. As fish use the main channel for movement, they could easily disperse mussels from one area to another (Woolnough 2006). Walleye commonly move 32-72 km in a year (Steuck 2004) and smallmouth bass may move up to 75 km during winter (Lyons & Kanehl 2002); therefore, if fish are carrying glochidia over the winter, they could become infested in one area and move considerable distances in the spring, effectively dispersing mussels. The combination of these habitat and life-history traits of mussels and their fish hosts could greatly facilitate long-distance dispersal for *L. cardium* in large rivers, leading to undifferentiated populations.

Alteration of habitat by impoundment of rivers may be fragmenting mussel populations in this system and changing how populations interact. Trends in this study that support evidence for fragmentation are statistically significant population differentiation that is greater in both pairwise F_{ST} values and the more sensitive exact tests at the pool level compared to the population level. This might indicate less gene flow through dams and more gene flow within a pool. Furthermore, assignment rates increased from 10 to 30% from the population to the pool level, which indicates less gene flow is now occurring across dams than historically occurred in the free-flowing

river. Trends may be weak in the current data because animals collected for this study were on average 27 years old (range 9-87 yrs; N = 190) based on age-length regression. Therefore not enough time has passed for a strong effect on population differentiation, because there is a time lag between when dams were constructed (about 80 years ago) and the time needed to register a difference in genetic diversity. Time lags are more pronounced in long lived animals such as mussels.

Habitat fragmentation for species with historically high gene flow is likely to have negative effects on both genetic diversity and population demography (Berg *et al.* 2008). Because dispersal and gene flow link populations, high genetic diversity is indicative of both high gene flow and large population sizes in this species. However, if linkages between populations are severed, then genetic diversity will be reduced, which in turn reduces population sizes, leading species into an extinction vortex (Ewers & Didham 2006; Frankham 2003; Templeton *et al*. 1990). Considering the imperiled status of 70% of all mussel fauna in North America, it is important to understand how population genetic processes are affected by anthropogenic activities. Results of this study along with those for several other large river mussel species suggest high gene flow and connectivity over rather large distances are quite common (Berg *et al*. 1998; Elderkin *et al*. 2007, 2008; chapters 2 and 3 of this thesis). If habitat fragmentation is occurring in large rivers, we may have a difficult time predicting extinction probabilities unless we have an understanding of genetic factors influencing population dynamics (Berg *et al*. 2008).

References

Amoros C, Bornette G (2002) Connectivity and biocomplexity in waterbodies of riverine floodplains. *Freshwater Biology,* **47**, 761-776.

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill AVS (2007) Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes,* **7**, 10-14.

Amyot JP, Downing JA (1997) Seasonal variation in vertical and horizontal movement of the freshwater bivalve *Elliptio complanata* (Mollusca:Unionidae). *Freshwater Biology,* **37**, 345-354.

Applied Biosystems (2008) GeneMapper® software version 4.0 reference and troubleshooting guide. Applied Biosystems, Foster City, CA.

Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, Massachusetts.

Avise JC (2004) *Molecular markers, natural history and evolution*. Sinauer Press, Sunderland, Massachusetts.

Becker GC (1983) *The Fishes of Wisconsin*. The University of Wisconsin Press, Madison, WI.

Berg DJ, Haag WR, Guttman SI, Sickel JB (1995) Mantle biopsy: a technique for nondestructive tissue-sampling of freshwater mussels. *Journal of the North American Benthological Society,* **14**, 577-581.

Berg DJ, Cantonwine EG, Hoeh WR, Guttman SI (1998) Genetic structure of *Quadrula quadrula* (Bivalvia:Unionidae): little variation across large distances. *Journal of Shellfish Research,* **17**, 1365-1373.

Berg DJ, Christian AD, Guttman SI (2007) Population genetic structure of three freshwater mussel (Unionidae) species within a small stream system: significant variation at local spatial scales. *Freshwater Biology,* **52**, 1427-1439.

Berg DJ, Levine TD, Stoeckel JA, Lang BK (2008) A conceptual model linking demography and population genetics of freshwater mussels. *Journal of the North American Benthological Society,* **27**, 395-408

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology,* **13**, 551-561.

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity,* **93**, 504-509.

Dawson KJ, Belkhir K (2001) A Bayesian approach to the identification of panmictic populations and the assignment of individuals. *Genetical Research,* **78**, 59-77.

Duran S, Pascual M, Estoup A, Turon X (2004a) Strong population structure in the marine sponge *Crambe crambe* (Poecilosclerida) as revealed by microsatellite markers. *Molecular Ecology,* **13**, 511-522.

Duran S, Pascual M, Turon X (2004b) Low levels of genetic variation in mtDNA sequences over the western Mediterranean and Atlantic range of the sponge *Crambe crambe* (Poecilosclerida). *Marine Biology,* **144**, 31-35.

Eackles MS, King TL (2002) Isolation and characterization of microsatellite loci in *Lampsilis abrupta* (Bivalvia:Unionidae) and cross-species amplification within the genus. *Molecular Ecology Notes,* **2**, 559-562.

Elderkin CL, Christian AD, Vaughn CC, Metcalfe-Smith JL, Berg DJ (2007) Population genetics of the freshwater mussel, *Amblema plicata* (Say 1817) (Bivalvia: Unionidae): evidence of high dispersal and post-glacial colonization. *Conservation Genetics,* **8**, 355-372.

Estoup A, Angers B (1998) Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In: *Advances in molecular ecology* (ed. Carvalho G), pp. 55-86. IOS Pres, Amsterdam, The Netherlands.

Ewers RM, Didham, RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews*, **81**, 117-142.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes -application to human mitochondrial-DNA restriction data. *Genetics,* **131**, 2, 479-491.

Fahrig WF (2003) Effects of habitat fragmentation on biodiversity. *Annual Reivew of Ecology, Evolution, and Systematics*, **34**, 487-515.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology,* **10**, 1500-1508.

Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies,* **326**, S22-S29.

Geist J, Kuehn R (2005) Genetic diversity and differentiation of central European freshwater pearl mussel (*Margaritifera margaritifera* L.) populations: implications for conservation and management. *Molecular Ecology,* **14**, 425-439.

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. *Journal of Heredity,* **86**, 485-486.

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

Haponski AE, Marth TA, Stepien CA (2007) Genetic divergence across a low-head dam: a preliminary analysis using logperch and greenside darters. *Journal of Great Lakes Research,* **33**, 117-126.

Holland-Bartels LE, Kammer TW (1989) Seasonal reproductive development of *Lampsilis cardium*, *Amblema plicata plicata*, and *Potamilus alatus* (Pelecypoda, Unionidae) in the Upper Mississippi River. *Journal of Freshwater Ecology,* **5**, 87-92.

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution,* **53**, 1898-1914.

Jager HI, Chandler JA, Lepla KB, Van Winkle W (2001) A theoretical study of river fragmentation by dams and its effects on white sturgeon populations. *Environmental Biology of Fishes,* **60**, 347-361.

Johnson RL, Liang FQ, Farris JL (1998) Genetic diversity among four Amblemini species (Bivalvia:Unionidae) in the Cache and White rivers, Arkansas. *Southwestern Naturalist,* **43**, 321-332.

Junk WJ, Bayley PB, Sparks RE (1989) The flood pulse concept in river-floodplain systems. *Archiv Fur Hydrobiologie*, **3**, 110-127.

Kelly MW, Rhymer JM (2005) Population genetic structure of a rare unionid (*Lampsilis cariosa*) in a recently glaciated landscape. *Conservation Genetics,* **6**, 789-802.

Kelner D (2003) Upper Mississippi River Mussel Species Accounts. U.S. Army Corps of Engineers, St. Paul, MN. PDF [Online] at: www.fws.gov/midwest/mussel/documents/umrs_mussel_distribution_kelner_2003.pdf.

King TL, Eackles MS, Gjetvaj B, Hoeh WR (1999) Intraspecific phylogeography of *Lasmigona subviridis* (Bivalvia : Unionidae): conservation implications of range discontinuity. *Molecular Ecology,* **8**, S65-S78.

Lydeard C, Cowie RH, Ponder WF, et al (2004) The global decline of nonmarine mollusks. *Bioscience,* **54**, 321-330.

Lyons J, Kanehl P (2002) *Seasonal movements of smallmouth bass in streams.* American Fisheries Society, Bethesda, Maryland.

Machordom A, Araujo R, Erpenbeck D, Ramos MA (2003) Phylogeography and conservation genetics of endangered European Margaritiferidae (Bivalvia: Unionoidea). *Biological Journal of the Linnaean Society,* **78**, 235-252.

McGregor SW, Garner JT (2004) Changes in the freshwater mussel (Bivalvia: Unionidae) fauna of the Bear Creek system of northwest Alabama and northeast Mississippi. *American Malacological Bulletin,***18**, 61-70.

Mulvey M, Lydeard C, Pyer DL, et al (1997) Conservation genetics of North American freshwater mussels *Amblema* and *Megalonaias*. *Conservation Biology,* **11**, 868-878.

Nagel KO, Badino G, Alessandria B (1996) Population genetics of European Anodontinae (Bivalvia: Unionidae). *Journal of Molluscan Studies,* **62**, 343-357.

National Native Mussel Conservation Committee (1998) National strategy for the conservation of native freshwater mussels. *Journal of Shellfish Research,* **17**, 1419-1428.

Natureserve (2008) NatureServe Explorer: an online encyclopedia of life [web application, http://www.natureserve.org/]. Version 7.0. **2008**, Accessed June 11, 2008.

Nilsson C, Reidy CA, Dynesius M, Revenga C (2005) Fragmentation and flow regulation of the world's large river systems. *Science,* **308**, 405-408.

Paetkau D, Slade R, Burden M, Estoup A (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 RDM, Holmes EC (1998) *Molecular evolution: a phylogenetic approach.* Blackwell Science Ltd., Malden, Massachussetts.

Parmalee PW, Bogan AE (1998) *The freshwater mussels of Tennessee*, first edition. The University of Tennessee Press, Knoxville,Tennessee.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* **6**, 288-295.

Pearse DE, Crandall KA (2004) Beyond F_{ST}: analysis of population genetic data for conservation. *Conservation Genetics,* **5**, 585-602.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology,* **12**, 844-855.

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity,* **95**, 536-539.

Pringle CM (2001) Hydrologic connectivity and the management of biological reserves: a global perspective. *Ecological Applications,* **11**, 981-998.

Quinn JW, Kwak TJ (2003) Fish assemblage changes in an Ozark river after impoundment: a long-term perspective. *Transactions of the American Fisheries Society,* **132**, 110-119.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America,* **94**, 9197-9201.

Raymond M, Rousset F (1995a) An exact test for population differentiation. *Evolution,* **49**, 1280-1283.

Raymond M, Rousset F (1995b) Genepop (Version-1.2) - population genetics software for exact tests and ecumenicism. *Journal of Heredity,* **86**, 248-249.

Reid SM, Wilson CC, Mandrak NE, Carl LM (2008) Population structure and genetic diversity of black redhorse (*Moxostoma duquesnei*) in a highly fragmented watershed. *Conservation Genetics,* **9**, 531-546.

Rice W (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.

Roe KJ, Lydeard C (1998a) Molecular systematics of the freshwater mussel genus P*otamilus* (Bivalvia: Unionidae). *Malacologia,* **39**, 195-205.

Roe KJ, Lydeard C (1998b) Species delineation and the identification of evolutionarily significant units: lessons from the freshwater mussel genus *Potamilus* (Bivalvia: Unionidae). *Journal of Shellfish Research,* **17**, 1359-1363.

Roe KJ, Hartfield PD, Lydeard C (2001) Phylogeographic analysis of the threatened and endangered superconglutinate-producing mussels of the genus *Lampsilis* (Bivalvia: Unionidae). *Molecular Ecology,* **10**, 2225-2234.

Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning, a laboratory manual.* Cold Spring Harbor Laboratory Press, New York.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics,* **152**, 1079-1089.

Schneider S, Roessli D, Excoffier L (2000) Arelquin, v. 2.0: a software for population genetics analysis. [Online] http://lgb.unige.ch/arlequin/software/.

Serb JM, Buhay JE, Lydeard C (2003) Molecular systematics of the North American freshwater bivalve genus *Quadrula* (Unionidae:Ambleminae) based on mitochondrial ND1 sequences. *Molecular Phylogenetics and Evolution,* **28**, 1-11.

Shaw PW, Pierce GJ, Boyle PR (1999) Subtle population structuring within a highly vagile marine invertebrate, the veined squid *Loligo forbesi*, demonstrated with microsatellite DNA markers. *Molecular Ecology,* **8**, 407-417.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.

Slatikin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139,** 457-462.

Stanford JA, Ward JV (2001) Revisiting the serial discontinuity concept. *Regulated Rivers: Research and Management,* **17**, 303-310.

Strayer DL (2006) Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society,* **25**, 271-287.

Steuck M (2004) Walleye and sauger on the upper Mississippi River. Unpublished abstract online at: www.iowadnr.com/fish/programs/research/mississ/wallumr.pdf.

Templeton AR, Shaw K, Routman E, Davis SK (1990) The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden,* **77**, 13-27.

van Oosterhout C, Weetman D, Hutchinson WF (2006) Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes,* **6**, 255-256.

Vaughn CC, Taylor CM (1999) Impoundments and the decline of freshwater mussels: a case study of an extinction gradient. *Conservation Biology*, **13**, 912-920.

Wang J (2004) Application of the one-migrant-per-generation rule to conservation and management. *Conservation Biology,***18**, 332-343.

Watters GT (1996) Small dams as barriers to freshwater mussels (Bivavlia, Unionoida) and their hosts. *Biological Conservation,* **75**, 79-85.

Watters GT, O'Dee SH (1999) Glochidia of the freshwater mussel *Lampsilis* overwintering on fish hosts. *Journal of Molluscan Studies,* **65**, 453-459.

Watters GT, O'Dee SH, Chordas S (2001) Patterns of vertical migration in freshwater mussels (Bivalvia: Unionoida). *Journal of Freshwater Ecology,* **16**, 541-549.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the analysis of population structure. *Evolution,* **38**, 1358-1370.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration, $F_{ST} \neq 1/(4Nm + 1)$. *Heredity*, **82**, 117-125.

Wiegand T, Moloney KA, Naves J, Knauer F (1999) Finding the missing link between landscape structure and population dynamics: a spatially explicit perspective. *The American Naturalist,* **154**, 605-627.

Wirth T, Bernatchez L (2001) Genetic evidence against panmixia in the European eel. *Nature,* **409**, 1037-1040.

Wlosinski JH, Marecek SR (1996) Fish movement on the Upper Mississippi River. *Project status report.* [Online] at: www.umesc.gov/reports_publications/psrs/psr_1996_04.html.

Woolnough, DA (2006) The importance of host fish in long range transport of unionids in large rivers. PhD thesis, Iowa State Univeristy, Ames, IA.

Wright S. (1978) *Evolution and the genetics of populations, volume 4, variation within and among natural populations*. University of Chicago Press, Chicago, IL.

Yamamoto S, Morita K, Koizumi I, Maekawa K (2004) Genetic differentiation of white-spotted charr (*Salvelinus leucomaenis*) populations after habitat fragmentation: spatial-temporal changes in gene frequencies. *Conservation Genetics,* **5**, 529-538.

Zanatta DT, Murphy RW (2006) Evolution of active host-attraction strategies in the freshwater mussel tribe *Lampsilini* (Bivalvia: Unionidae). *Molecular Phylogenetics and Evolution,* **41**, 195-208.

Zanatta DT, Murphy RW (2007) Range-wide population genetic analysis of the endangered northern riffleshell mussel, *Epioblasma torulosa rangiana* (Bivalvia: Unionoida). *Conservation Genetics,* **8**, 1393-1404.

Zanatta DT, Murphy RW (2008) The phylogeographical and management implications of genetic population structure in the imperiled snuffbox mussel, *Epioblasma triquetra* (Bivalvia: Unionidae). *Biological Journal of the Linnaean Society,* **93**, 371-384.

Zanatta DT, Fraley SJ, Murphy RW (2007) Population structure and mantle display polymorphisms in the wavy-rayed lampmussel, *Lampsilis fasciola* (Bivalvia: Unionidae). *Canadian Journal of Zoology-Revue Canadienne De Zoologie,* **85**, 1169-1181.

Table 1. Sample size (N), number of alleles with number of private alleles in parentheses (Na), allelic richness (Ar), observed (H_o) and expected (H_E) heterozygosity, for *Lampsilis cardium* sampled in 6 pools of the upper Mississippi River, Boone River, IA, and Lake St. Clair, MI.

			Locus								
Pool	Population		111	10	206	67	213	2	24	29	mean
4	1	N	23	23	23	23	23	23	22	22	22.8
		Na	14	21 (1)	16	8	20	8	2	24 (1)	11.3
		Ar	9.4	12.4	11.0	5.5	12.0	5.4	2.0	13.9	9.0
		Ho	0.61*	0.65*	0.52*	0.57	0.87	0.91	0.36	0.91*	0.7
		He	0.91	0.94	0.94	0.74	0.94	0.76	0.51	0.97	0.84
7	2	N	20	20	20	20	20	20	20	19	19.9
		Na	14	22	12	8 (2)	19	7	3	20	13.9
		Ar	9.9	13.8	8.8	6.6	12.3	5.1	2.5	13.3	9.0
		Ho	0.90	0.75*	0.50*	0.70	0.95	0.85	0.35	0.84*	0.7
		He	0.91	0.97	0.89	0.80	0.95	0.76	0.48	0.96	0.84
7	3	N	34	35	35	35	35	35	35	33	34.6
		Na	17	28	19	8	25	8	4 (1)	25	18.6
		Ar	10.7	13.7	11.0	6.0	13.1	4.9	2.7	13.3	9.4
		Ho	0.76*	0.71*	0.69*	0.57*	0.94	0.69*	0.49	0.85*	0.7
		He	0.93	0.97	0.93	0.83	0.96	0.70	0.39	0.96	0.83
7	4	N	19	20	20	20	20	20	20	20	19.9
		Na	17	16	17	8	22	8	3	22	14.1
		Ar	11.4	10.9	10.7	5.7	12.7	5.5	2.5	13.1	9.1
		Ho	0.79*	0.50*	0.50*	0.60	0.80*	0.65	0.60	0.95	0.7
		He	0.94	0.93	0.92	0.74	0.94	0.76	0.53	0.96	0.84
8	5	N	23	25	24	26	25	25	25	22	24.4
		Na	17 (1)	20	19	7	21	6 (1)	3	22	15.3
		Ar	10.6	12.1	11.4	5.7	12.5	4.7	2.6	12.6	9.0
		Ho	0.87	0.76	0.71*	0.62	0.84*	1.00	0.52	0.68*	0.7
		He	0.92	0.95	0.93	0.73	0.95	0.74	0.53	0.95	0.84
8	6	N	16	16	17	17	17	17	17	16	16.6
		Na	15	18	14	5	19	6	2	18	12.1
		Ar	10.9	12.9	10.5	3.4	12.9	4.6	2.0	13.2	8.8
		Ho	0.81	0.69*	0.76*	0.24	0.88	0.94	0.41	0.69*	0.7
		He	0.93	0.96	0.92	0.27	0.96	0.70	0.45	0.96	0.77
8	7	N	21	18	20	21	21	21	20	18	20.0
		Na	16	17	17	7	19	5	3	19 (1)	12.0
		Ar	10.7	12.1	11.6	4.5	12.0	4.1	2.5	12.9	8.8
		Ho	0.90	0.61*	0.70*	0.33*	0.62*	0.81	0.45	0.61*	0.6
		He	0.92	0.95	0.94	0.53	0.95	0.70	0.44	0.96	0.80

Table 1 (continued)

8	8	N	20	20	20	20	20	19	20	19	19.8
		Na	12	16	16	9	21	8	2	19	12.9
		Ar	9.3	11.5	10.5	6.3	12.8	5.6	2.0	12.6	8.8
		Ho	0.85	0.45*	0.70*	0.65	1.00	0.95	0.60	0.79	0.7
		He	0.91	0.94	0.92	0.74	0.95	0.76	0.47	0.95	0.83
8	9	N	20	20	20	20	20	19	20	19	19.8
		Na	12	16	16	9	21	8	2	19	12.9
		Ar	9.3	11.5	10.5	6.3	12.8	5.6	2.0	12.6	8.8
		Ho	0.85	0.45*	0.70*	0.65	1.00	0.95	0.60	0.79*	0.7
		He	0.91	0.94	0.92	0.74	0.95	0.76	0.47	0.95	0.83
8	10	N	11	11	10	11	11	9	11	10	10.5
		Na	12	15	10	6	15	4	2	14	9.8
		Ar	11.0	13.1	9.4	5.9	13.2	4.0	2.0	13.1	9.0
		Ho	1.00	0.82	0.70	0.27*	0.82	0.78	0.64	0.80*	0.7
		He	0.94	0.96	0.89	0.84	0.97	0.69	0.51	0.96	0.85
9	11	N	20	20	20	20	20	20	19	19	19.8
		Na	14	19	12	8	18	6	2	22	12.6
		Ar	10.5	12.5	9.3	5.8	11.3	5.0	2.0	13.7	8.8
		Ho	0.80*	0.75*	0.70*	0.60	0.75*	0.80	0.47	0.74*	0.7
		He	0.93	0.95	0.91	0.65	0.92	0.76	0.51	0.97	0.83
9	12	N	20	20	20	20	20	20	20	20	20.0
		Na	16	19	11 (1)	6	22	4	2	22	13.0
		Ar	11.7	12.7	8.9	5.5	12.8	3.7	2.0	13.6	8.9
		Ho	0.85	0.70*	0.45*	0.80	0.90	1.00	0.35	0.80*	0.7
		He	0.95	0.96	0.90	0.71	0.95	0.69	0.41	0.97	0.82
9	13	N	20	20	20	20	20	20	20	20	20.0
		Na	13	20	15	8	19 (1)	6	2	25	12.7
		Ar	9.8	12.6	10.1	5.4	12.8	4.6	2.0	14.5	9.0
		Ho	0.80*	0.65*	0.50*	0.55	0.90	0.85	0.40	0.95	0.7
		He	0.92	0.95	0.91	0.62	0.96	0.74	0.43	0.97	0.81
9	14	N	20	20	20	20	20	20	20	20	20.0
		Na	15	16	14	8	19 (1)	6	2	19	11.4
		Ar	11.0	11.5	10.1	6.0	12.3	4.6	2.0	11.9	8.7
		Ho	0.95	0.85	0.65*	0.65*	0.95	0.85	0.50	0.95*	0.8
		He	0.94	0.94	0.92	0.72	0.95	0.69	0.43	0.94	0.82
11	15	N	24	24	24	24	24	24	23	24	23.9
		Na	15	19	13	8	21 (1)	6	3	21	12.1
		Ar	10.8	12.1	9.9	6.0	12.2	4.7	2.4	12.5	8.8
		Ho	0.96	0.58*	0.79*	0.67*	0.83*	0.88	0.39	0.79*	0.7
		He	0.94	0.95	0.92	0.73	0.95	0.74	0.43	0.95	0.83

Table 1 (continued)

14	16	N	26	25	26	25	26	26	25	25	25.5
		Na	17	23	16	7	25	9	2	29 (1)	14.1
		Ar	11.3	12.7	10.5	5.1	13.2	5.7	2.0	14.6	9.4
		Ho	0.81	0.64*	0.54*	0.48	0.85*	0.88	0.48	0.76	0.7
		He	0.94	0.95	0.92	0.66	0.96	0.74	0.47	0.97	0.83
NA	17	N	25	25	24	25	25	25	24	25	24.8
	(Boone)	Na	14 (1)	25	16 (1)	6	21	7	2	23 (1)	12.2
		Ar	8.4	10.4	10.9	6.1	11.1	6.1	4.3	11.9	8.7
		Ho	0.72	0.68	0.71	0.64	0.92	0.84	0.25	0.96	0.7
		He	0.90	0.97	0.94	0.56	0.96	0.74	0.28	0.96	0.79
NA	18	N	16	20	20	20	19	20	20	18	19.1
	(Lake St. Clair)	Na	10	15 (1)	14	8	17 (1)	10 (1)	17 (4)	17	12.3
		Ar	11.0	13.3	10.8	5.8	12.7	5.0	2.4	13.5	9.3
		Ho	0.56	0.60	0.80	0.40	0.89	0.60	0.50	0.67	0.6
		He	0.89	0.92	0.94	0.79	0.93	0.73	0.67	0.95	0.85

*Indicates locus-by-population combinations with H_o significantly different from H_e, after a sequential Bonferroni correction (Rice 1989), experiment-wide $\alpha = 0.05$.

Table 2. Pairwise estimates of population differentiation for *Lampsilis cardium* in the UMR only. F_{ST} values are below the diagonal and the number of significant locus-by-population pair exact tests from the genic differentiation option in Genepop (Raymond & Rousset 1995b) are above the diagonal.

Pop.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		1	4*	2*	2	1	2	4*	4*	2	2	2	4	2	3*	1
2	0.006		1	2	0	1*	1	1	1	1	1	0	2	1	1	1
3	0.017	0.005		3	1	3*	1	3*	3*	1	2	0	1	1	1	3*
4	0.012	0.006	0.016		1	2	2	1	1	2	1	0	1	1	1	3
5	-0.001	-0.001	0.011	-0.002		1	1	2	2	0	0	0	0	1	0	1
6	0.016	0.027	0.035*	0.018	0.006		0	2	2	2*	1	0	1	0	2	1
7	0.010	0.011	0.015	0.008	0.000	0.001		2	2	2	0	0	0	0	0	1
8	0.010	0.007	0.019*	0.002	0.001	0.014	0.003		0	2	2	0	2*	3	1	2
9	0.010	0.007	0.019	0.002	0.001	0.014	0.003	-0.026		2	2	0	2*	3	1	2
10	0.005	0.005	0.015	0.011	0.010	0.046*	0.026	0.019	0.019		3	1	3	2	2	0
11	0.007	0.014	0.025	0.002	0.000	0.012	0.002	0.005	0.005	0.015		0	3	0	0	1
12	0.009	0.000	0.004	-0.004	-0.005	0.004	0.000	-0.001	-0.001	0.016	0.002		0	1	0	0
13	0.012	0.009	0.012	0.004	-0.002	0.003	-0.005	0.010	0.010	0.022	0.006	-0.007		1	0	1
14	0.007	0.003	0.007	-0.002	-0.003	0.006	-0.001	0.008	0.008	0.018	0.002	-0.006	0.000		2	1
15	0.011	0.004	0.012	0.001	-0.003	0.012	-0.001	-0.002	-0.002	0.017	0.002	-0.007	0.000	0.001		0
16	0.002	0.003	0.019*	0.008	-0.003	0.010	0.003	0.003	0.003	0.011	0.007	0.002	0.004	0.005	-0.001	

*Indicates significance after adjusting for multiple comparisons via a sequential Bonferroni correction (Rice 1989), experiment wide $\alpha = 0.05$.

Table 3. Pairwise estimates of population differentiation for *Lampsilis cardium* in the upper Mississippi River, Boone River, IA, and Lake St. Clair, MI. F_{ST} values are below the diagonal and the number of significant locus-by-population pair exact tests from the genic differentiation option in Genepop (Raymond & Rousset 1995b) are above the diagonal. Populations are grouped into pools from the upper Mississippi River, labels follow Figure 1.

Pool	4	7	8	9	11	14	Boone
4		4*	4*	4*	3*	3	-
7	0.009		4*	2	1	3*	-
8	0.007	0.007*		4*	1	3*	-
9	0.009	0.005	0.002		0	3	-
11	0.014	0.006	0.001	0.0002		2*	-
14	0.002	0.008	0.003	0.006	0.002		-
Boone	0.021*	0.015*	0.009	0.006	0.005	0.015	
Lake St. Clair	0.011	0.017*	0.024*	0.024*	0.025*	0.011	0.041*

*Indicates significance after adjusting for multiple comparisons via a sequential Bonferroni correction (Rice 1989), experiment wide $\alpha = 0.05$.

Table 4. AMOVA results for *Lampsilis cardium*. Panel a for tests at the population and pool levels within the UMR only. Panel b for tests at the largest spatial scale including the populations in the Mississippi River, Boone River, IA, and Lake St. Clair, MI. F-statistics are Weir and Cockerham (1984) estimates.

Panel a

Source of variation	d.f.	ss	% variation	P	F-statistics
Among pools	5	24	0.3	0.112	$\theta = F_{CT} = 0.003$
Among populations within pools	10	38	0.5	0.122	$F_{SC} = 0.004$
Within populations	670	2137	99.2	0.006	$F_{ST} = 0.003$
Total	685	2199	100.0		

Panel b

Source of variation	d.f.	ss	% variation	P	F-statistics
Among drainages	2	14	1.1	0.04	$\theta = F_{CT} = 0.011$
Among populations within drainage	5	24	0.5	0.00	$F_{SC} = 0.005$
Within populations	766	2440	98.4	0.00	$F_{ST} = 0.016$
Total	773	2478	100.0		

Table 5. Bayesian assignment results (Rannala & Mountain 1997) for each individual into populations and pools that were sampled for *Lampsilis cardium* as described in the text. Bold face numbers indicate the number of individuals correctly assigned to each population or pool.

	Mississippi River																	L. St. Clair
	Pool 4	Pool 7			Pool 8						Pool 9				Pool 11	Pool 14	Boone	
Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	**4**	-	1	3	1	-	1	-	-	-	1	-	-	-	-	2	2	2
2	-	**1**	3	5	2	-	1	-	-	-	-	1	1	1	1	-	1	2
3	2	3	**6**	9	2	-	5	-	-	2	2	4	3	5	3	2	3	-
4	1	-	-	**-**	1	1	1	-	-	1	1	2	3	3	1	1	1	1
5	3	6	1	5	**6**	5	6	-	-	-	4	3	2	2	4	6	6	2
6	1	-	-	-	1	**1**	1	-	-	-	2	2	1	-	-	4	1	-
7	1	1	3	2	1	1	**-**	-	-	2	-	-	2	1	1	1	1	1
8	1	2	-	-	1	-	-	**-**	20	-	3	1	-	-	1	-	1	1
9	1	-	-	-	1	-	1	19	**-**	-	-	-	-	-	-	-	-	2
10	1	1	1	-	1	-	1	-	-	**1**	1	1	-	-	1	3	1	1
11	1	1	-	4	-	-	-	-	-	-	**2**	-	-	-	-	2	-	1
12	-	1	-	-	1	2	-	1	1	-	-	**1**	-	3	2	1	-	-
13	-	1	1	-	-	-	-	-	-	-	1	-	**2**	-	-	1	2	-
14	1	-	-	1	-	1	-	-	-	-	-	-	3	**2**	2	-	-	-
15	-	-	-	1	-	-	-	-	-	1	2	1	1	1	**2**	-	1	-
16	7	1	4	3	9	5	4	-	-	4	4	3	5	3	6	**2**	6	3
17	-	2	-	1	1	1	-	-	-	1	1	1	-	-	-	1	**1**	-
18	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1	-	**6**
% correct	17	5	30	0	23	6	0	0	0	9	10	5	10	10	8	8	4	30
Pool																		
Pool 4	**4**	-	1	3	1	-	1	-	-	-	1	-	-	-	-	2	2	2
Pool 7	3	**4**	**9**	**14**	5	1	7	-	-	3	3	7	7	9	5	3	5	3
Pool 8	8	10	5	7	**11**	**7**	**9**	**19**	**20**	**3**	10	7	5	3	7	14	10	7
Pool 9	2	3	1	5	1	3	-	1	1	-	3	1	5	5	4	4	2	1
Pool 11	-	-	-	1	-	-	-	-	-	1	**2**	**1**	**1**	**1**	2	-	1	-
Pool 14	7	1	4	3	9	5	4	-	-	4	4	3	5	3	**6**	2	6	3
Boone	-	2	-	1	1	1	-	-	-	1	1	1	-	-	-	**1**	**1**	-
Lake St. Clair	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1	-	**6**
% correct	17	20	45	40	42	41	43	95	100	27	10	5	5	5	25	4	4	30

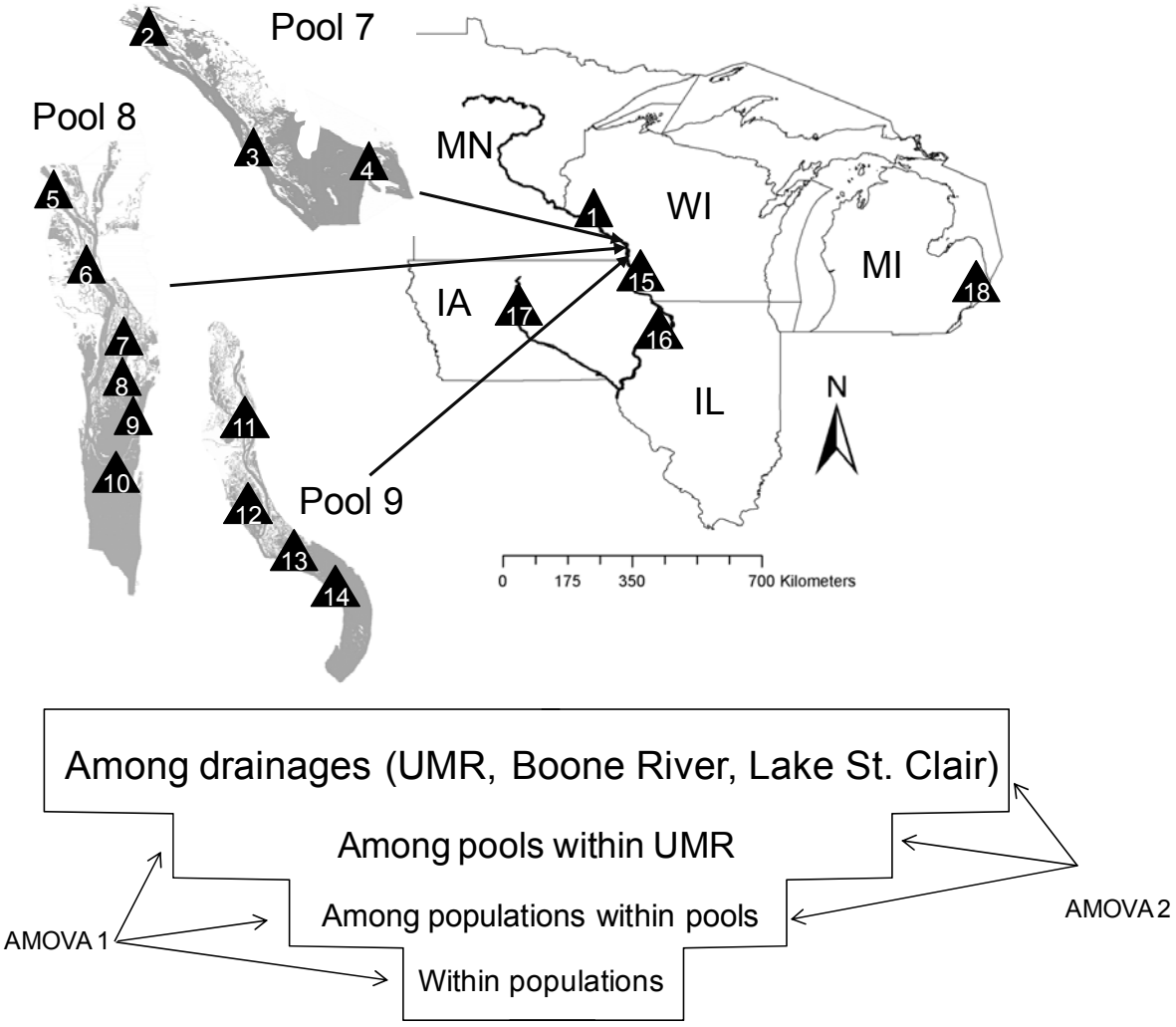

Figure 1. Map of populations where *Lampsilis cardium* were sampled in the upper Mississippi River (UMR), Boone River, IA, and Lake St. Clair, MI. UMR populations are numbered 1-16, Boone River is number 17, Lake St. Clair is number18. Pools with multiple populations sampled are enlarged (not to scale) for detail, grey indicates water, and white indicates land. Hierarchy of variation is depicted in the inverted pyramid, and the levels included in each AMOVA (Excoffier *et al*. 1992) as indicated.

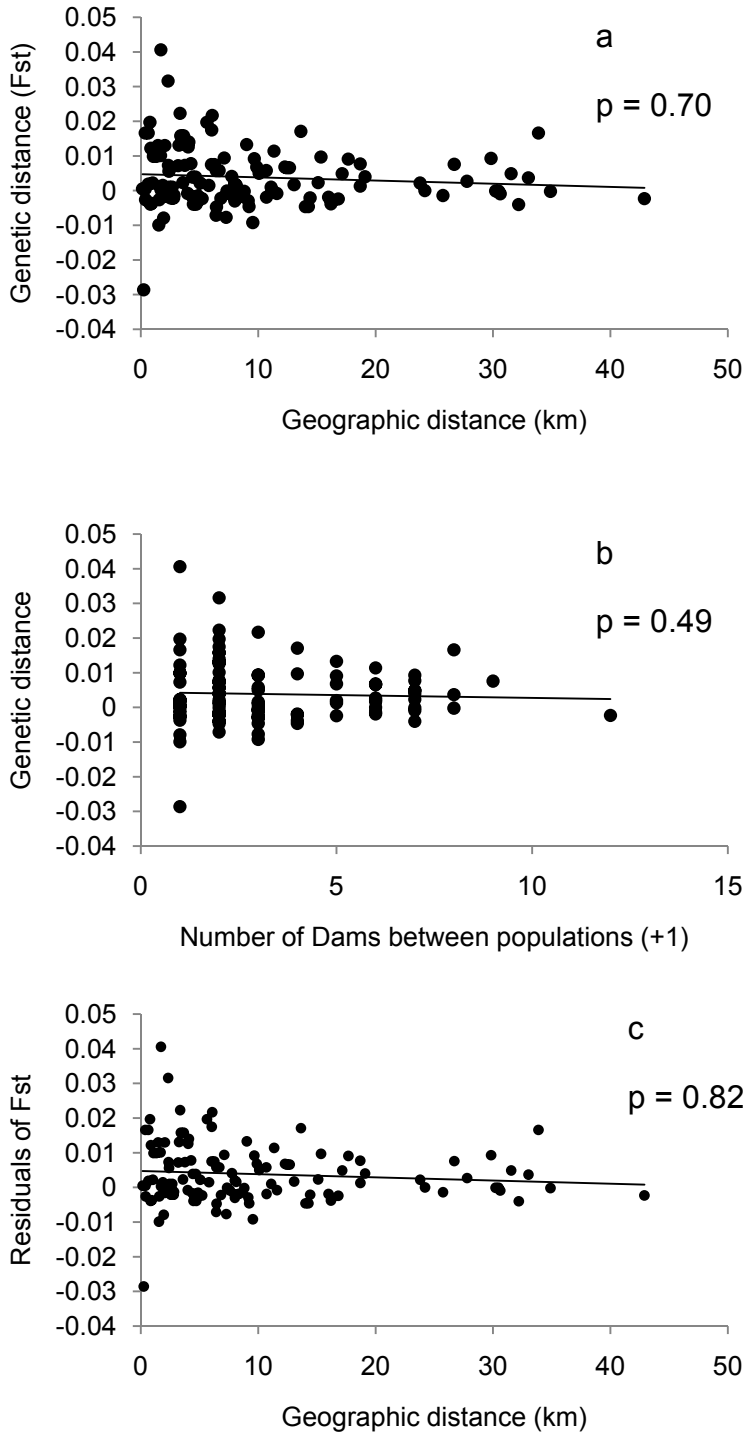

Figure 2. Correlations of (a) geographic distance and genetic distance, (b) the number of dams between populations, and (c) geographic distance and residuals of pairwise F_{ST} values.

Chapter 5

Conclusions

North America has the most diverse assemblages of freshwater mussels in the world (Bogan 1993), but they are also one of the most imperiled faunal groups on the continent (Williams *et al.* 1993). Molecular genetic studies of freshwater mussels provide information that is useful in guiding conservation of this fauna (Berg *et al.* 2008; National Native Mussel Conservation Committee 1998). The results of my work show that large-scale patterns of genetic diversity in one common (*Lampsilis cardium*) and one rare (*Cumberlandia. Monodonta*), species of freshwater mussel are quite similar, with high levels of genetic diversity but relatively low population genetic structure.

The natural hierarchy of river systems provided an ideal framework for evaluating genetic diversity in mussels at various spatial scales. Genetic diversity was quantified among individuals within populations, among populations within a river reach, among rivers, and in some cases, among drainage systems. Both species I studied are very widespread, but one is common and the other is highly imperiled. Nonetheless, both species had the greatest genetic diversity at the smallest spatial scale (within populations), and substantially less genetic diversity was contained at the largest spatial scales (among rivers or among drainage basins). The lack of strong population structure may be indicative of their large-river habitat and the fact that they rely on fish for part of their life cycle. The Flood-Pulse concept (Junk *et al.* 1989) of river systems may explain dispersal over greater distances than may be expected. This concept suggests that large flood-plain rivers act as super-highways for fish, promoting movement among habitats that provide resources such as food, nesting or spawning sites, and overwintering habitat. As fish use the main stems of large rivers to move among these areas, they could easily disperse mussels from one river reach to another. Prior research indicates that over 50% of the variation in mussel communities is explained by the distribution and abundance of host fish (Vaughn & Taylor 2000). The importance of host fish in structuring mussel distribution in the upper Mississippi River was recently described (Woolnough 2006).

Since fish and mussels are so tightly coupled in nature, one would expect mussels to have similar patterns in genetic diversity. Phylogeographic patterns have

been found to be similar among fish species in North America (Mayden 1988; Wiley & Mayden 1985). For example, populations in the north have less genetic diversity as a result of their recent founding post-Pleistocene compared to the longer stable histories of southern populations (Bernatchez & Wilson 1998). This has also been found in several other mussel species (Elderkin *et al.* 2008; Zanatta & Murphy 2008), but, northern populations of *C. monodonta* and *L. cardium* did not follow this common trend. Another common finding is the evidence for multiple glacial refugia (Bernatchez 2001). *Lampsilis cardium* populations in Lake St. Clair had evidence of colonization from multiple refugia, and so did northern populations of *C. monodonta*. Explicit hypothesis testing with statistical phylogeography indicated that the Driftless Area of the upper midwest was not a likely refugium for the *C. monodonta* population in the St. Croix River. For *C. monodonta*, Pleistocene glaciations created some intraspecific diversity, but post-Pleistocene dispersal was effective in homogenizing these populations and eliminating most of the structure among populations. Similarly, for *L. cardium*, post-Pleistocene dispersal and high gene flow is more prominent in shaping current genetic structure among populations than isolation events during the Pleistocene.

High levels of gene flow were inferred for both species studied. Unfortunately, rivers across the globe have been fragmented by impoundments (Nilsson *et al.* 2005). For species with a normally high level of gene flow, isolation is likely to have negative effects on both genetic diversity and population demography (Frankham 2003). For example, if gene flow between populations is reduced, then genetic diversity will be reduced. This in turn can reduce population sizes and put a species into an extinction vortex (Ewers & Didham 2006; Frankham 2003). The species I studied are at different places along the continuum of an extinction vortex. *Lampsilis cardium* is still relatively common and widespread with many populations in most of its former range. As indicated by the fine-scale study in the Mississippi River, there has not been a strong reduction in overall genetic diversity, and gene flow remains relatively high. On the other hand, *C. monodonta* has experienced drastic population losses and only survives as a few populations at the edges of its former range. Clearly, the lack of intervening populations has effectively isolated these populations and gene flow has been reduced

to near zero levels. This creates very different conservation scenarios for these species.

The results of this study for *C. monodonta* indicate that remaining populations would make good sources for recovery efforts involving translocations, with a few caveats. First, reintroductions would need to be done with enough individuals to maintain current levels of genetic diversity (Lydeard & Roe 1998). Unfortunately, how many individuals are needed to meet these goals is elusive. Furthermore, given that this species was formerly widespread and had connectivity among populations, any reintroduced populations would be threatened by fragmentation of rivers. Since the host is unknown, reestablishment of populations alone may not be sufficient for recovery. Two situations might develop: (1) establishment of new populations will restore connectivity throughout the range of the species because hosts will be able to move up and down the rivers; these rivers won't be fragmented for the purposes of *C. monodonta*; (2) establishment of new populations that remain isolated from other populations because hosts will still encounter barriers to dispersal — this will leave these populations vulnerable to extirpation and also lead to drifting apart of populations over evolutionary time unless continuing management efforts are pursued to connect populations.

The results for *L. cardium* indicate as long as fish remain connected to mussel communities, *L. cardium* may be able to maintain the same level of gene flow it has had since the glaciers receded 8,000-10,000 years ago. The Mississippi River is managed with locks and dams, which only allow for limited fish movement, so increasing fish passage can only help maintain mussel dispersal and high levels of genetic diversity. It is also important to manage habitats that support mussel communities, maintain zebra mussel-free refuges, and control some of the other negative impacts on mussels such as sedimentation (Lydeard *et al.*2004). In rivers that have low-head dams without mechanisms that allow fish passage, it has been shown that dams as low as one meter can stop fish dispersal (Watters 1996). To maintain gene flow for mussels in these systems, fish movement must be considered in management or recovery plans. Conservation plans in these systems would need to actively maintain connections among mussel populations.

References

Berg DJ, Levine TD, Stoeckel JA, Lang BK (2008) A conceptual model linking demography and population genetics of freshwater mussels. *Journal of the North American Benthological Society,* **27**, 395-408.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Neararctic and Palearctic fishes. *Molecular Ecology,* **7**, 431-452.

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

Bogan AE (1993) Freshwater bivalve extinctions (Mollusca, Unionoida) - a search for causes. *American Zoologist,* **33**, 599-609.

Elderkin CL, Christian AD, Metcalfe-Smith JL, Berg DJ (2008) Population genetics and phylogeography of freshwater mussels in North America, *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia : Unionidae). *Molecular Ecology,* **17**, 2149-2163.

Ewers RM, Didham RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews,* **81**, 117-142.

Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies,* **326**, S22-S29.

Junk WJ, Bayley PB, Sparks RE (1989) The flood pulse concept in river-floodplain systems. *Archiv Fur Hydrobiologie*, **3**, 110-127.

Lydeard C, Roe KJ (1998) Phylogenetic systematics: the missing ingredient in the conservation of freshwater unionid bivalves. *Fisheries*, **23**, 16-17.

Lydeard C, Cowie RH, Ponder WF et al. (2004) The global decline of nonmarine molluscs. *Bioscience*, **54**, 321-330.

Mayden RL (1988) Vicariance biogeography, parsimony, and evolution in North American freshwater fish. *Systematic Zoology*, **37**, 329-355.

National Native Mussel Conservation Committee (1998) National strategy for the conservation of native freshwater mussels. *Journal of Shellfish Research,* **17**, 1419-1428.

Nilsson C, Reidy CA, Dynesius M, Revenga C (2005) Fragmentation and flow regulation of the world's large river systems. *Science,* **308**, 5720, 405-408.

Strayer DL (2006) Challenges for freshwater invertebrate conservation. *Journal of the North American Benthological Society,* **25**, 271-287.

Vaughn CC, Taylor CM (2000) Macroecology of a host-parasite relationship. *Ecography,* **23**, 11-20.

Watters GT (1996) Small dams as barriers to freshwater mussels (Bivalvia, Unionoida) and their hosts. *Biological Conservation,* **75**, 79-85.

Wiley EO, Mayden RL (1985) Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. *Annals of the Missouri Botanical Garden,* **72,** 596-635.

Williams JD, Warren ML, Cummings KS, Harris JL, Neves RJ (1993) Conservation status of freshwater mussels of the United States and Canada. *Fisheries,* **18**, 6-22.

Woolnough, DA (2006) The importance of host fish in long range transport of unionids in large rivers. PhD thesis, Iowa State University, Ames, IA.

Zanatta DT, Murphy RW (2008) The phylogeographical and management implications of genetic population structure in the imperiled snuffbox mussel, *Epioblasma triquetra* (Bivalvia: Unionidae). *Biological Journal of the Linnaean Society,* **93**, 371-384.