GENETIC IDENTIFICATION AND PHYLOGENETICS OF LAKE WACCAMAW ENDEMIC FRESHWATER MUSSEL SPECIES

Kristine Sommer

A Thesis Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
of the Requirements for the Degree of
Master of Science

Department of Biology and Marine Biology

University of North Carolina Wilmington

2007

Approved by

Advisory Committee

Dr. Ami E. Wilbur

Dr. Arthur Bogan

Dr. Wilson Freshwater

Dr. Michael McCartney
Chair

Accepted by

_____
Dean, Graduate School

TABLE OF CONTENTS

ABSTRACT

Approximately 70% of modern-day freshwater mussel species in North America are considered threatened, endangered, or recently extinct, and a large number of the non-extinct species are endemic to a narrow geographic range. Freshwater mussel conservation efforts have been limited by taxonomic ambiguity and morphologic convergence. Lake Waccamaw in southeastern North Carolina contains two endemic species, *Lampsilis fullerkati* and *Elliptio waccamawensis*, which share nearly identical shell morphologies. This convergence in shell morphology complicates conservation efforts. To provide an alternative means to discriminate them, I developed a Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) assay for genetic identification. Genomic DNA was obtained using a non-lethal method of hemolymph extraction. DNA was amplified using 16S rRNA gene specific primers and digested with *Hinf* I, *Ava* II, and *Hind* III. However, only the banding patterns of *Ava* II and *Hind* III digestions were diagnostic for these species and were used to type 112 individuals. RFLP and DNA sequencing data revealed three individuals that had been misidentified based on morphology. In addition, phylogenetic analysis was used to assess the taxonomy and to test the status of these putative endemics. Mitochondrial 16S rRNA gene, *cox1*, and *nad1* sequences were obtained from 109 individuals sampled from Lake Waccamaw, the adjacent Waccamaw River, and the Yadkin/Pee Dee, Little Pee Dee, and Lumber Rivers in the Pee Dee Drainage. Results from Bayesian analyses suggest the endemic status of both *L. fullerkati* and *E. waccamawensis* may need to be reconsidered. *L. fullerkati* is not phylogenetically distinct from *Lampsilis radiata* collected outside the lake, and *E. waccamawensis* groups with and is not genetically distinguishable from *E. congaraea* individuals from the Waccamaw River.

ACKNOWLEDGEMENTS

I would like to thank Dr. Michael McCartney for all his guidance and advice. He was a great advisor and a great teacher in class and in the lab. He was always there to answer my questions and help in every way that he could. In addition to teaching me all there is to know about molecular biology, he also taught me how to be a better scientist and researcher.

I would also like to thank my committee members, Dr. Ami Wilbur, Dr. Wilson Freshwater, and Dr. Arthur Bogan for their advice and expertise. I would additionally like to thank Dr. Bogan for the topotype specimen samples, which were a significant contribution to my data.

In addition, I would like to thank Dr. Ryan Heise and Rob Nichols at the North Carolina Wildlife Resource Commission and Karen Lynch and her crew at the North Carolina Department of Transportation for their assistance in the field with collecting and identifying the animals. I would also like to thank the NCDOT for funding. Without the help of everyone above, this thesis would not have been possible. Thank you.

DEDICATION

I would like to dedicate this master's thesis to my husband Ben; without his support, I would not have made it this far. He has always encouraged me and stood by my side. I would also like to dedicate this to my family. Without my family, I would not be the person I am today.

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: Literature review and background information

Characteristics of freshwater mussels

Freshwater mussels are members of the Phylum Mollusca, Class Bivalvia, Subclass Palaeoheterodonta, and Order Unionoida (Campbell 2000; Roe and Hoeh 2003; Graf and Cummings 2006). Within Bivalvia, freshwater mussels are related to members of the subclass Heterodonta, which contains clams and other freshwater bivalves such as the Asian clam (Corbiculidae) and the zebra mussel (Dreissenidae) (Campbell 2000; Bieler and Mikkelsen 2006).

Similar to other bivalves, freshwater mussels possess a muscular foot, a visceral mass, mantle tissue, and the calcareous shell, which is secreted from the mantle tissue (McMahon and Bogan 2001; Brusca and Brusca 2003). Freshwater mussels, like other bivalves, are suspension feeders; however, unionid freshwater mussels do not have true siphons. Instead, they use apertures, formed by the extension of the mantle edges, to draw in and expel water (McMahon and Bogan 2001). Once the water has been drawn into the animal, it is passed over a set of ctenidia on which food particles are collected, sorted, and transported to the mouth (McMahon and Bogan 2001; Brusca and Brusca 2003). In addition to feeding, the ctenidia function in respiration and circulation, acting as a surface for gas exchange. Freshwater mussels maintain an open circulatory system, in which hemolymph bathes the tissues and organs, providing them with the oxygen needed for metabolism (McMahon and Bogan 2001). Because of their feeding behavior, freshwater mussels are primarily sedentary, spending most of their time partially buried just below the sediment surface.

The traits that set freshwater mussels apart from other bivalves and molluscs are features of their life history. In most bivalves, gametes are released freely into the water column, where fertilization occurs. The embryos develop into a trochophore larva, which further develops into a

veliger larva (Brusca and Brusca 2003). In contrast, the majority of freshwater mussel unionid species brood their embryos within their gills. This varies from using all four gills to only outer gills or a restricted portion of the outer gills in Unionidae. The embryos develop into a unique larva called the glochidium (Kat 1984; McMahon and Bogan 2001). Glochidia are ectoparasites of fish. Fish gills and fins provide a protective and nutrient rich environment in which the larvae can grow and develop. Because glochidia must attach to a host fish to metamorphose into juveniles, several freshwater mussels have adapted ways to increase the frequency of larvae encountering a proper fish host (Parmalee and Bogan 1998). For example, some species have developed mantle extensions, or lures (Parmalee and Bogan 1998; Bogan 1998), and others have developed superconglutinates (Bogan 1998; Roe et al. 2001), both are used to attract fish. In addition to adaptations of the adults, some unionids have developed glochidia with hooked valves allowing them to attach more efficiently to their hosts (Kat 1984). Freshwater mussels rely on specific host fishes for the development and survival of their larvae; this relationship plays an important role in freshwater mussel adaptation and speciation (Kat 1984).

Conservation of freshwater mussels

North America contains the greatest diversity of freshwater mussels, globally (Williams et al. 1993). However, approximately 70% of modern-day freshwater mussels have been assigned the status of recently extinct, endangered, threatened, or of special concern (Williams et al. 1993; Bogan 1998; Turgeon et al. 1998). Several factors are responsible for the drastic decline in North American freshwater mussel diversity, most of which are due to human impacts (Williams et al. 1993; Bogan 1998). Agricultural and industrial activities, such as mining, have decreased water quality by increasing sedimentation and pollution, respectively (Williams et al. 1993; Bogan 1998). Habitat loss as a result of dam building, dredging, and channel creation is

another threat to freshwater mussels (Williams et al. 1993; Bogan 1998). Artificial channels and dams disrupt unionid habitats by altering water flow and increasing siltation (Fuller 1974). Furthermore, dams reduce the reproductive success of freshwater mussels by disrupting mussel gametogenesis and by disrupting the natural environment of fish, which perform a primary role in the development of glochidia larvae (Fuller 1974; Williams et al. 1993).

In addition to habitat destruction, freshwater mussel populations are affected by commercial exploitation. Commercial harvest of freshwater mussels for both pearls and the manufacture of pearl buttons occurred during the late 19th and the 20th centuries (Parmalee and Bogan 1998). The button industry declined and ended by the mid 1960's because technology allowed for the cheaper and more efficient production of plastic buttons. However, exploitation of freshwater mussels for the cultured pearl industry increased where the button industry left off (Parmalee and Bogan 1998). In the 1990s, the harvest of freshwater mussels for the pearl industry increased, and only federally listed species were protected (Williams et al. 1993).

An additional threat to freshwater mussels is the spread of invasive species, such as the Asian clam, *Corbicula fluminea*, and the zebra mussel, *Dreissena polymorpha* (Williams et al. 1993; Bogan 1998; Lydeard et al. 2004). Invasive species take over in great numbers and stress freshwater mussel populations that are already declining (Williams et al. 1993). For example, *C. fluminea* grows in high densities in areas of suitable habitat, causing reductions in native populations of freshwater mussels (Belanger et al. 1990). The reduction of unionid populations in areas of high Asian clam populations could be due to competition for food (Leff et al. 1990). Zebra mussels also lower unionid fitness by impeding locomotion and burrowing, feeding, and reproduction (Haag et al. 1993; Tucker 1994). The elimination of local populations of native

freshwater mussels by invasive species is of particular concern in areas that contain endangered or rare species (Ricciardi et al. 1998).

Taxonomy of freshwater mussels

There are two superfamilies of freshwater mussels: Unionoidea and Etherioidea. Unionoidea contains the families Unionidae, Margaritiferidae, and Hyriidae; Etherioidea contains the families Etheriidae, Iridinidae, and Mycetopodidae (Hoeh et al. 1998a, 2001; Roe and Hoeh 2003; Bogan 2004). Unionoida is a globally diverse order, absent only from Antarctica (Hoeh et al. 1998a; Roe and Hoeh 2003). The vast majority of species within Unionoida are members of the family Unionidae, with North America hosting approximately 43% of Unionoidan diversity (Graf and Cummings 2006). Unionidae in North America contains 50 genera, which are divided into 278 species and 13 subspecies (Turgeon et al. 1998; McMahon and Bogan 2001; Roe and Hartfield 2005). Of the North American unionids, the majority are from the subfamily Ambleminae (Campbell et al. 2005). Ambleminae is further divided into several tribes including Pleurobemini and Lampsilini, which contain the genera *Elliptio* and *Lampsilis*, respectively (Campbell et al. 2005). These genera are the focus of this thesis project.

The majority of taxonomic information on freshwater mussels is based on shell morphology and internal anatomy (McMahon and Bogan 2001; Campbell et al. 2005). Some examples of shell characteristics used to classify freshwater mussels are shell shape, presence or absence of rays on the outer shell surface, shell sculpture present or absent, and presence or absence of hinge teeth inside the shell (McMahon and Bogan 2001; Serb et al. 2003; Campbell et al. 2005). Internal anatomy, such as gill and corresponding reproductive structures have been used to classify freshwater mussels (Serb et al. 2003; Campbell et al. 2005). However, conchological classification has often been ambiguous. One main source of confusion results

from convergence and plasticity of shell phenotype, which allows freshwater mussels to adapt to their surrounding environment (McMahon and Bogan 2001). Therefore, heterospecific mussels that live in the same environment are likely to share similar shell characteristics and at the same time conspecifics that live in different regions could adapt shell characteristics specific to their environments (Serb et al. 2003).

In addition to the convergence of conchological characters used to classify freshwater mussels, inconsistent use of names during early classification confounds unionid taxonomy (Johnson 1970; Campbell et al. 2005). Furthermore, consistent collections of freshwater mussels from some regions were not made until the early 1900's (Johnson 1970). Johnson (1970) reviewed the historical classification of Unionidae from the Atlantic slope region; after which, he provided an updated report of the systematics of this region, taking care to be conservative in recognizing species.

It is crucial for researchers to define the taxonomy and ecology of freshwater mussels because so many of them are highly threatened (Fuller 1977). Williams et al. (1993) produced a report on the status of freshwater mussels from the United States and Canada in order to provide agencies with important information for the management and conservation of freshwater mussels. Turgeon et al. (1998) also published a list of both scientific and common names of North American freshwater and marine molluscs, including unionid mussels, in addition to their status. The authors provided a consistent list using both types of nomenclature, with the goal of eliminating further taxonomic confusion (Turgeon et al. 1998). Correct taxonomy of freshwater mussels is important in identifying and providing reliable censuses of endangered of threatened species in areas of development and construction. In addition, knowing the correct taxonomy and distributions of freshwater mussels will help ensure successful management efforts, such as

reintroductions for restoration of depleted local populations. It is important therefore, that systematic research of freshwater mussels continues alongside management efforts.

Biogeography of the southeastern Atlantic slope

The geographic region of this study is the southern Atlantic slope, which extends from the James River in Virginia to the Altamaha River in Georgia (Johnson 1970). It is separated from the interior basin rivers to the west by the Appalachian Mountains (Johnson 1970; Sepkoski and Rex 1974). The Appalachian Mountains were formed during the Paleozoic Era (Hack 1969; Johnson 1970). During the Cenozoic Era, the Coastal Plain was shaped by plate movements and by sea level changes induced by the formation and melting of glaciers during the late Pliocene/Pleistocene (Hack 1969). Analysis of coastal plain sediments has demonstrated that this region was both above and below sea level at different times during this era (Hack 1969). There is some evidence for stream capture events in this region, in which one stream - often having a smaller drainage area - merges with and becomes "captured" into the watershed of another stream (Hack 1969). Van der Schalie (1945) provided examples of cases in which distribution patterns of freshwater mussels could be used as evidence for historical patterns of stream confluence and denounced the theory that aquatic birds served as a mode of dispersal of freshwater mussels. By solidifying the fact that freshwater mussels were primarily dispersed by their host fish, Van der Schalie (1945) demonstrated that freshwater mussel distributions can serve as a valuable tool for understanding the histories of stream flow patterns because of their poor ability to disperse among drainages (Van der Schalie 1945; Johnson 1970). Johnson (1970) speculated that there was evidence of stream capture events in his investigation of the distribution of freshwater mussels in the Atlantic slope region. Johnson (1970) believed that some of the southern Atlantic slope species descended from western species, which were

transported to the southern Atlantic slope during the confluence of the Savannah River with Apalachicola River from the west, prior to the Pleistocene Epoch. However, the southern Atlantic slope also contains species that are endemic to the southern Atlantic slope in addition to species that are endemic to individual drainages (Johnson 1970).

This project occurs in North Carolina where there are eleven major rivers, the Savannah, Yadkin/Pee Dee, Lumber, Waccamaw, Cape Fear, White Oak, Neuse, Pamlico, Roanoke, Chowan, and Pasquotank (Bogan 2002). Samples for this study were taken from the Yadkin/Pee Dee, Lumber, and Waccamaw River systems. In addition, two samples were taken from the Tar-Pamlico and Neuse River basins. Samples were collected from four genera of freshwater mussels – *Lampsilis*, *Leptodea*, *Elliptio*, and *Uniomerus*; however, the study mainly focuses on two of the genera, *Elliptio* and *Lampsilis*. These genera were chosen because of their endemic species found in Lake Waccamaw, North Carolina.

Biogeography and endemism

A new understanding of the geographic distribution and the formation of species came about with McArthur and Wilson's theory of island biogeography (1963, 1967). Island biogeography is centered on two main principles; area geographic effects on species diversity and equilibrium rate of immigration and extinction (McArthur and Wilson 1963, 1967). Under equilibrium conditions, in which rates of immigration are equal to rates of extinction, islands that are larger and closer to the mainland will contain a greater number of species than islands that are smaller and farther away. McArthur and Wilson (1967) also refer to habitat islands, which are patches of habitat on a continental landmass that are surrounded by dissimilar habitat types. In 1974, Sepkoski and Rex applied multiple regression analyses to coastal river populations of freshwater mussels to determine if coastal rivers were habitat islands that analogously act as

oceanic islands. The authors determined that coastal rivers of the Atlantic slope, Gulf slope, and peninsular Florida do act as geographic islands and that area of the river, serving as a source of colonists, was the primary factor affecting species diversity. Rivers also acted as stepping stones and rivers farther away from the source population contained fewer species than closer rivers. Also, small rivers had fewer species than larger rivers (Sepkoski and Rex 1974). Furthermore, Browne (1981) determined that freshwater lakes also act as oceanic islands; larger lakes contained a greater number of species than smaller lakes.

Island biogeography is useful in studying endemic species, which are those that are restricted to a geographic area (Cox and Moore 1993). Areas that have been geographically isolated longer are more likely to have endemic species than more recently isolated areas (Cox and Moore 1993). In fact, Sepkoski and Rex (1974) concluded that the great abundance of endemic species in southern Atlantic slope and eastern Gulf slope coastal rivers was due to the fact that the southern rivers were less glaciated and had more time for endemic species to form. In addition to age of isolation, stability of the environment also contributes to the degree of endemism present; the greater the stability, the higher the frequency of endemic species (Cox and Moore 1993). Also, island size can determine the amount of endemic species present. Smaller islands are unable to form as many endemic species as larger islands because smaller islands are more prone to extinction and have higher turnover rates than larger islands (Mayr 1965).

Endemic species are of primary concern for conservation biologists because of their relation to biodiversity. Kerr (1997) measured species richness and endemism in various taxa and found that within a taxon, endemism is significantly correlated with species richness; areas with greater numbers of species had a greater fraction of endemics. Because of this correlation,

endemic species are often used by conservation biologists to identify areas of concern or "biodiversity hotspots" (Myers et al. 2000; Myers 2003). Hotspots are defined as areas that contain a large number of endemic species found in threatened habitats (Myers et al. 2000; Myers 2003). By using hotspots, conservation biologists can more effectively protect biodiversity within the limits of funding (Myers et al. 2000; Myers 2003). In relation to freshwater mussels, Lydeard et al. (2004) suggested that the assignment of biodiversity hotspots for nonmarine molluscs should be used to enhance management efforts of these rapidly declining animals.

In addition to being indicators of biodiversity, endemic species are also of conservation concern because of the effect of range restriction on their genetic diversity. Frankham (1998) proposed that endemic species were more inbred and prone to extinction than nonendemic species due to the fact that endemic species are older and have had more time for inbreeding depression to accumulate. Frankham (1998) compared inbreeding coefficients for several studies and determined that both endemic and nonendemic island populations were more inbred than mainland populations. Further, among island populations, endemic populations were even more inbred than nonendemic populations (Frankham 1998). Because of their range restriction, endemics can have small population sizes, which make them more prone to genetic drift and inbreeding, thus resulting in a loss of genetic diversity (Frankham et al. 2002). This loss of genetic diversity poses an even greater risk to highly threatened species, such as freshwater mussels.

There are various methods for quantifying endemism. Kerr (1997) did so by giving endemism scores to a wide range of taxa by summing the inverse of the number of quadrats in which each species was found. Alternatively, Peterson and Navarro-Siguenza (1999) and Evans

et al. (2003) measured endemism based on the presence of monophyly; that is, the finding that endemic species consist of groups of individuals that are the exclusive descendents of a single common ancestor. In the present study, the presence of monophyly will be used to confirm or discount the endemic status of two freshwater mussel species found in Lake Waccamaw.

Lake Waccamaw endemics

Lake Waccamaw is the largest Carolina bay lake and is located in Columbus County, North Carolina (Porter and Horn 1980; Casterlin et al. 1984). Lake Waccamaw differs from other Carolina bay lakes due to its neutral pH (Casterlin et al. 1984; Frey 1951), higher total productivity, drainage patterns, and occurrence of endemism (Frey 1951). Three endemic species of fish reside solely in Lake Waccamaw: *Fundulus waccamensis*, *Etheostoma perlongum*, and *Menidia extensa* (Hubbs and Raney 1946). Two additional species have been added to the list of possible endemics, a madtom (Jenkins and Palmer 1978) and a pygmy sunfish (Shute et al. 1981). In addition, Lake Waccamaw is the home of endemic gastropods and bivalves (Fuller 1977). The gastropods are the Waccamaw Snail, *Amnicola sp*. and the Waccamaw Scavenger Snail, *Lioplax subcarinata*. The endemic freshwater mussels (the subjects of the present study) are the Waccamaw Spike, *Elliptio waccamawensis* (Lea 1863) and the Waccamaw Fatmucket, *Lampsilis fullerkati* (Johnson 1984). In addition to the endemic species, several other freshwater mussels are commonly found in Lake Waccamaw, such as *Elliptio fisheriana*, *Toxolasma pullus*, *Lampsilis* sp., *Lampsilis crocata*, and *Leptodea ochracea* (Porter and Horn 1980; Porter and Horn 1983). The amount of endemism present in Lake Waccamaw is interesting because the lake is geologically young. The age of the lake is estimated to be around 15,000 years old (Stager and Cahoon 1987). It is therefore surprising that

so much speciation has apparently occurred over this short of time in a temperate zone ecosystem (Hubbs and Raney 1946).

The Lake Waccamaw endemic species, *L. fullerkati* and *E. waccamawensis* are the focus of this thesis project. Lake Waccamaw provides a unique system for the development of molecular markers for species identification and for testing the utility of non-lethal hemolymph sampling as a source of DNA. Even though *E. waccamawensis* and *L. fullerkati* are listed as state endangered and threatened, respectively (Bogan 2002), they are locally abundant within the lake (Porter and Horn 1980, 1983). However, *E. waccamawensis* is much more abundant than *L. fullerkati* (Bogan 2002; Porter and Horn 1983). In addition, even though they are members of different genera, *L. fullerkati* and *E. waccamawensis* are morphologically very similar (Bogan 2002). Therefore, the development of a Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) assay for distinguishing these two species will be important for assessing population health and size because any misidentifications of *L. fullerkati* could produce significant error due to this species' lower abundance. Further, the development of this PCR-RFLP could be applied to other systems beside Lake Waccamaw. There are other species within these genera that are difficult to distinguish based on shell morphology alone. In fact, *Elliptio* and *Lampsilis* individuals collected from local North Carolina rivers were often confused (Chapter Three of the present study). For example, my phylogenetic analyses demonstrated that *Elliptio complanata* individuals share similar shell morphologies with *Uniomerus carolinianus* individuals from the same rivers. In addition, there was evidence that *Lampsilis radiata radiata* were morphologically similar to a cryptic, undescribed species of *Lampsilis*.

Goals of this project

With the use of Restriction Fragment Length Polymorphisms (RFLPs) I plan to identify the Lake Waccamaw endemic species *Lampsilis fullerkati* and *Elliptio waccamawensis,* based on the presence or absence of restriction enzyme recognition sequences. These two species are nearly identical in shell morphology and are often confused in the field. Therefore, the development of genetic markers distinguishing these two species from each other would greatly facilitate their identification. I hope to use this as a first step towards developing a diagnostic key based on RFLPs as a quick method of species identification, which can be used by field biologists to supplement morphologic identification. In addition, by sampling individuals of *Lampsilis* and *Elliptio* from sites along the Pee Dee, Waccamaw, and Lumber River drainages, I hope to determine if the Lake Waccamaw endemic species are truly endemic, based on the presence of monophyly. I hope to gain some insight into where the endemic species came from and to which species from the surrounding drainages they are most closely related.

CHAPTER TWO: Genetic identification of *Lampsilis fullerkati* and *Elliptio waccamawensis* using PCR-RFLP

INRODUCTION

The greatest diversity of freshwater mussels occurs in North America (Williams et al. 1993). However, 70% of freshwater mussels are considered endangered, extinct, or threatened (Williams et al. 1993). The life histories of freshwater mussels make them highly susceptible to anthropogenic impacts and environmental alteration (Williams et al. 1993; Bogan 1998). However, conservation efforts are hindered by poor taxonomy and phenotypic plasticity in conchological features (Campbell et al. 2005). In addition, because of their highly threatened status, research in the past was limited to widespread, healthy populations (Berg et al. 1995).

Over the past decade or so, some non-lethal methods have been developed for the study of freshwater mussels. Berg et al. (1995) demonstrated through field experiments of two freshwater mussel species that removing a 1 cm^2 piece of mantle tissue did not cause significant mortality in treatment animals versus control animals. The use of mantle snips was employed in several genetic studies since the development of the technique (Henley et al. 2006). However, Henley et al. (2006) found evidence that taking mantle snips was an invasive process and that it could potentially lead to mortality. The authors presented a less invasive method, which involved brush swabbing the foot and viscera of *Quadrula pustulosa* individuals. They determined that extracted DNA concentrations from swab samples were comparable to those of mantle snips, and that extracted swab DNA was successful in amplification and sequencing (Henley et al. 2006). Further, Gustafson et al. (2005) tested the effects of taking hemolymph samples from the anterior adductor muscle of *Elliptio complanata* individuals. These individuals were collected from the field and kept in laboratory conditions to study the effects of hemolymph extraction on the growth and survival of these animals. The authors conducted two studies, one which tested the short term effects of sampling, and the other which tested for the effects of

repeatedly sampling over a period of seven months (Gustafson et al. 2005). From both of their studies they determined that the removal of 0.5 ml of hemolymph from the anterior adductor muscle had no significant effects on growth and survival. The authors concluded that hemolymph extraction was a non-lethal method that could be used for the continual health monitoring of freshwater mussels (Gustafson et al. 2005). Raley et al. (2006) expanded on the Gustafson et al. (2005) study by demonstrating that DNA extracted from hemolymph of *E. complanata* individuals was just as reliable for genetic analyses as DNA extracted from mantle tissue.

In spring of 2004, McCartney and Wilbur (2007) tested the lethality of hemolymph extraction from the anterior adductor muscle of freshwater mussels in the field. They set up two enclosures in Lake Waccamaw, North Carolina; each contained 20 *Elliptio waccamawensis,* 16 *Lampsilis fullerkati,* and 20 *Leptodea ochracea* individuals. Hemolymph was extracted from half of the animals in each enclosure, the other half were left as controls. The animals were tagged and monitored for two months after hemolymph extraction (McCartney and Wilbur 2007). Similar to Gustafson et al. (2005), McCartney and Wilbur (2007) found no significant effect of hemolymph extraction on the growth and survival of these freshwater mussel species, this time, under conditions in the field. The next step, which was the focus of this thesis, was to test the usefulness of the collected hemolymph as a source of genomic DNA for genetic identification of freshwater mussels.

Genetic markers have been used as a tool for studying freshwater mussel systematics and populations since the early 1980s. Pioneer studies which combined the use of genetic techniques and morphology to describe freshwater mussels paved the way for further research (Davis and Fuller 1981; Davis et al. 1981; Davis 1984; Kat 1983a, 1983b, 1983c; Kat and Davis 1984).

Davis and Fuller (1981) and Davis (1984) highlighted the need for the use of genetic markers in determining species status of freshwater mussels, which was previously defined by shell morphology. The main reservations the authors had with relying on shell morphologies was that there were not enough distinctive characters to classify species and that the shell characters which did exist exhibited evidence of phenotypic convergence (Davis and Fuller 1981; Davis 1984). These studies employed immunoelectrophoretic (Davis and Fuller 1981) and allozyme (Davis et al. 1981; Davis 1984; Kat 1983a,b,c; Kat and Davis 1984) techniques to measure genetic variation at different taxonomic levels for a variety of freshwater mussels. Through these studies, the authors gained a better understanding of the taxonomic relationships and species richness of the groups they were studying (Davis 1984).

These early studies thoroughly demonstrated the value of using genetic markers in freshwater mussel taxonomy and population studies. Throughout the next two decades, a variety of molecular markers have been implemented in freshwater mussel research. Additional studies with allozymes were applied to populations of *Lampsilis* from North Carolina (Stiven and Alderman 1992), *Margaritifera hembli* in Louisiana (Curole et al. 2004), and of *Velesunio* species in Australia (Baker et al. 2003; Hughes et al. 2004). In addition to allozymes, Hughes et al. (2004) utilized mitochondrial DNA sequences of the cytochrome oxidase subunit I gene (*cox1*). Buhay et al. (2002) also used mitochondrial genes to measure genetic variation in populations of endangered species of freshwater mussels. Krebs (2004) utilized both the maternal and paternal copies of the 16S rRNA gene mitochondrial gene to study populations of the species *Pyganodon grandis*, in which mitochondrial DNA, as in other freshwater mussel genera, is inherited through both maternal and paternal parents. Microsatellite loci were used to

assess population genetic structure of *Lampsilis abrupta* (Eackles and King 2005) and *Lampsilis cariosa* (Kelly and Rhymer 2005); both of which are of conservation concern.

Another molecular technique that has been used with freshwater mussels is restriction fragment length polymorphism (RFLP) (White et al. 1994; 1996). RFLPs utilize restriction enzymes, which are found in bacteria. These enzymes cleave DNA at specific sequences, which are known as the enzyme's recognition sequence (Madigan et al. 2003; Avise 2004). DNA that is digested with these enzymes will be cut into predictable fragments, depending on which enzyme is used and where the recognition sequences lies on the gene; the resulting fragments can be visualized using gel electrophoresis (Avise 2004). A review of the methods of RFLPs is provided by Dowling et al. (1996).

White et al. (1994) developed a method using Polymerase Chain Reaction (PCR) and RFLPs to identify freshwater mussel glochidia on host fishes. This method would allow researchers to identify the species of freshwater mussels found on specific host fishes (White et al. 1994). They developed a PCR-RFLP analysis that would recognize and digest glochidia PCR products without contamination from the host fish DNA (White et al. 1994). The authors determined the nuclear ITS-1 region was specific for amplifying glochidia DNA. ITS-1 PCR products digested with the enzymes *Msp* I, *Bam* HI, and *Sau* 96I could be used to distinguish most of the freshwater mussel species that they were studying; however, they did find enzymes which failed to distinguish between species (White et al. 1994). White et al. (1996) continued their study and developed a diagnostic PCR-RFLP key that could be used to distinguish among glochidia found in French Creek, Pennsylvania. The key consisted of steps of amplification and digestion with restriction enzymes. The key works in a hierarchical fashion, like a morphological identification key; species are identified based the banding patterns produced

using a specified gene and restriction enzyme (White et al. 1996). The authors confirmed the accuracy of their PCR-RFLP key by following through it with tissue extracts from adults that were previously identified based on morphology (White et al. 1996). Overall, White et al. (1996) demonstrated that the use of PCR-RFLP for identification of freshwater mussel glochidia was inexpensive and did not require extensive training. In addition they suggested that this method would be beneficial in studying adult freshwater mussels (White et al. 1996).

In the present study, I developed a PCR-RFLP method using hemolymph DNA for the identification of the Lake Waccamaw endemics species *Lampsilis fullerkati* and *Elliptio waccamawensis*. These species share nearly identical shell morphologies; male *L. fullerkati* individuals are distinguished from *E. waccamawensis* by slight differences, such as the sharpness of the posterior ridge (Bogan 2002). The development of a PCR-RFLP technique using hemolymph extraction could aid field biologists in monitoring freshwater mussel populations and would especially be useful for species of conservation concern because of the non-lethal sampling technique.

MATERIALS AND METHODS

Sample collection

Animals were collected from Lake Waccamaw in Columbus County, North Carolina (Figure 1). In April 2004, 40 *Elliptio waccamawensis*, 40 *Leptodea ochracea*, and 32 *Lampsilis fullerkati* individuals were transferred to enclosures for a caging experiment at site 1 (Figure 1) to test the mortality effects of hemolymph extraction (McCartney and Wilbur 2007). These individuals were collected after the experiment was completed.

In August of 2004, 30 meter long transect lines were drawn at sites 2 and 3 (Figure 1). Animals were collected along the length of the transect lines within 1 meter to the left or the

right of the line. The number of individuals sampled from each transect depended on their availability. With the aid of freshwater biologists from the North Carolina Wildlife Resource Commission (NCWRC) and the North Carolina Department of Transportation (NCDOT), individuals were sorted by species. Once the animals were sorted, 20-100 µl of hemolymph was extracted from the anterior adductor muscle of a subsample of animals, using a 1 ml sterile syringe with a 27G1/2 needle. The animals were then measured with a caliper, photographed, and returned to the location from where they were obtained.

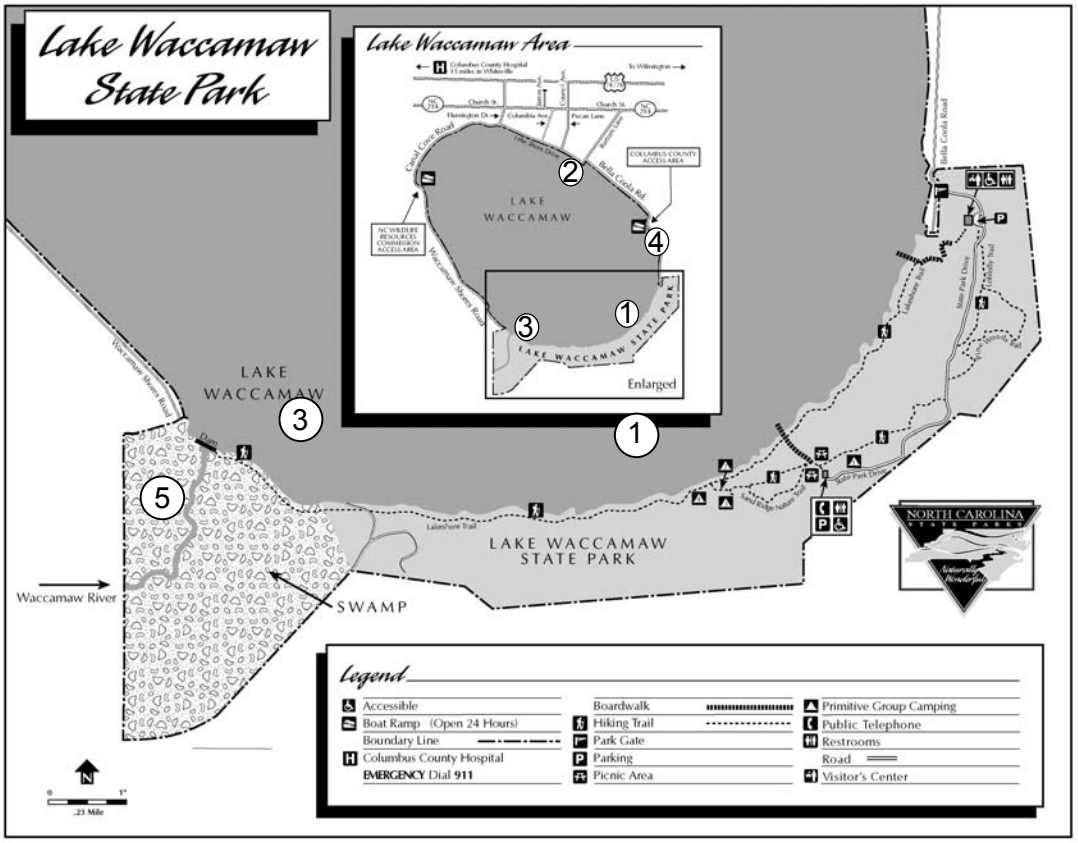

Figure 1. Map displaying Lake Waccamaw sampling locations. Site 1 represents the location of the field enclosures from the spring 2004 mortality experiments. Sites 2 and 3 are the locations of the August 2004 transect lines. Site 4 is from where the *L. radiata radiata* Big Creek sample was collected in 2001 (Chapter 3). Site 5 is location from which some of the 2005 Waccamaw River samples were taken (Chapter 3). The map was obtained from Lake Waccamaw State Park.

DNA extraction

DNA was extracted from tissue of the control animals from the Lake Waccamaw enclosures using a PureGene DNA extraction kit (Gentra Systems, Minneapolis MN). The kit procedure and reagents were modified to accommodate small tissue volumes. A total of 200 µl of cell lysis solution and 1 µl of Proteinase K solution were combined with a 2-8mg piece of tissue. The samples were incubated overnight at 55ºC. Proteins, RNAs, and other cellular materials were separated from the DNA using 70 µl of protein precipitation solution and centrifugation. The DNA was then precipitated with 100% isopropanol and washed with 70% ethanol. Once the ethanol was dried off, the DNA was resuspended in 35 µl sterile water. DNA samples were stored in a minus 20ºC freezer to prevent degradation. DNA was extracted from hemolymph using the same protocol as for tissue, except that 1.5 µl of Proteinase K solution and 50 µl of hemolymph were combined with the cell lysis solution for a total of 200 µl to begin the extraction.

Purification of DNA extracts

A StrataPrep PCR Purification Kit (Stratagene, La Jolla CA) was used to purify the DNA extracts through centrifugation through a silica gel affinity column. The DNA was bound to the column, washed, and eluted with 50 µl of sterile water. For hemolymph extracts, after elution the samples were dried down and re-suspended in 10-15 µl of sterile water to obtain greater amplification success. Both purified tissue DNA and hemolymph DNA extracts were stored at minus 20ºC. For extended storage hemolymph and body tissues were stored at minus 40ºC.

PCR amplification of 16S rRNA gene

A portion of the *16s ribosomal RNA (rRNA)* region of the mitochondrial genome was amplified using universal *16s* primers (Lydeard et al. 1996): 16sARLMyt (5`

CGACTGTTTAACAAAAACAT 3`) and 16sBRHMyt (3` ACATGTGCTGAGTTCAGAACGG 5`). Internal PCR and sequencing primers were designed from our initial freshwater mussel DNA sequences to improve success with these species: 16sUN693F (5'AGATAATGCCTGCCCAGTG 3') and 16sUN1178R (5' CGGTCTTAACTCAGCTCGTGTA 3'). PCR reactions used 1X PCR buffer with 1.5 mM $MgCl_2$ [Applied Biosystems (ABI), Foster City, CA], 0.2 mM each dNTP, 0.5 µM each primer, and 1 U *Taq* polymerase (ABI) in a 25 µl final volume. Cycling parameters were: an initial 5:00 at 94ºC, followed by 35 cycles of (94ºC for 1:00, 50ºC for 1:00, 72ºC for 2:00), followed by a final 5:00 soak at 72ºC. Reactions were carried out using a PTC-100 Thermal Cycler (MJ Research Inc., Waltham MA).

Purification of PCR products

Primers and salts were removed from the amplified segments of DNA using the StrataPrep PCR Purification Kit. The procedure was the same as listed above for the purification of the DNA extracts. PCR products that were used for restriction fragment length polymorphism (RFLP) were dried down and re-concentrated in 25 µl sterile water, which resulted in stronger bands during gel electrophoresis, and provided template for DNA sequencing when necessary.

Restriction Fragment Length Polymorphism (RFLP)

Purified PCR products from *L. fullerkati* and *E. waccamawensis* were digested with *Hinf* I, *Ava* II, and *Hind* III restriction endonucleases (New England BioLabs, Beverly MA), which were chosen as candidate diagnostic restriction enzymes based on internal DNA sequencing (Figure 3). All digests were completed in 20 µl volume reactions, containing 10 µl of the following cocktails and 10 µl of purified PCR Products. *Ava* II cocktails contained 0.5 x Buffer #4 (New England BioLabs), 8 units *Ava* II (10 units/µl), and sterile H_2O. *Hinf* I cocktails contained 0.5 x Buffer #2 (New England BioLabs), 8 units *Hinf* I (10 units/µl), and sterile H_2O.

Hind III cocktails contained 0.5 x Buffer #2, 20 units *Hind* III (20 units/µl), and sterile H_2O. The samples were incubated at 37ºC overnight, for at least 16 hours. The following day, the digested products were mixed with 4 µl of 6 x loading dye and loaded onto 1.8% NuSieve 3:1 agarose gels (Cambrex Bio Science Rockland, Inc., Rockland ME) containing Ethidium Bromide (0.25 µg/ml). Below, Table 1 shows the predicted RFLP banding patterns.

Table 1. Predicted 16S rRNA gene RFLP banding patterns for *L. fullerkati* and *E. waccamawensis*. The size of the fragments, in base pairs, cut by each enzyme is displayed for *L. fullerkati* and *E. waccamawensis* individuals.

Restriction Enzyme	Species	Fragment Size (base pairs)
Hinf I	*L. fullerkati*	440
	E. waccamawensis	238, 202
Ava II	*L. fullerkati*	280, 160
	E. waccamawensis	440
Hind III	*L. fullerkati*	440
	E. waccamawensis	120, 320

DNA sequencing

DNA sequences were produced to confirm or reject any discrepancies between morphologic and RFLP identifications. Purified PCR products were sequenced in the forward and reverse direction using 0.33 μM primers 16sUN693F and 16sUN1178R, respectively. Sequencing reactions were completed using Big Dye version 3.1 kits (ABI) and were loaded onto an ABI 3100 Genetic Analyzer. Sequences were edited using *Sequencher* (Gene Codes Corporation, Ann Arbor MI), organized in *MacClade 4.0* (Maddison and Maddison 2000), and aligned using *Clustal X* (Thompson et al. 1994). After the alignment was created, the file was imported back into *MacClade* and saved as a nexus file for analysis.

Initially, DNA sequences were produced to determine the reliability of the restriction enzymes *Hinf* I and *Ava* II. Sequences from *L. fullerkati* and *E. waccamawensis* were aligned and the recognition sequences of the restriction enzymes were checked for any polymorphisms, which would cause the diagnostic RFLP to not be fixed across all individuals within the species (Figure 3). The recognition sequences for *Hinf* I and *Ava* II are 5' GANTC 3' and 5' GG(A/T)CC 3,' respectively. Polymorphisms were found in *E. waccamawensis* individuals at the *Hinf* I recognition site. No polymorphisms were found in *L. fullerkati* individuals for either enzyme (Figure 3). *Hind* III was selected as an alternative to *Hinf* I. The recognition sequence for this enzyme was 5' AAGCTT 3.' No polymorphisms were identified within either species at this recognition site (Figure 3). Therefore, *Ava* II and *Hind* III were chosen as the diagnostic enzymes for the PCR-RFLP and were applied to all of the Lake Waccamaw samples.

Phylogenetic analysis

A maximum likelihood analysis was performed in *PAUP 4.0b10* (Swofford 2002) to efficiently check discrepancies between morphology and RFLP results at the DNA sequence

level. The best fit model of molecular evolution was determined by *ModelTest version 3.06* (Posada and Crandall 1998) to be TrN + I, a Tamura-Nei (Tamura and Nei 1993) with a proportion of invariable sites (0.6294). The model commands were appended to the alignment file and implemented in the maximum likelihood analysis. The phylogenetic tree was created using a maximum likelihood heuristic search with 10 random sequence additions and TBR branch swapping. Bootstrap analyses were performed using a heuristic search (10 random additions, TBR) and 100 replicates.

RESULTS

Utility of hemolymph as a source of DNA

Of the Lake Waccamaw enclosure animals, DNA taken from tissue of the control animals was 100% successful for both *L. fullerkati* and *E. waccamawensis* (Table 2A). Hemolymph DNA was successfully amplified for 93.8% of *L. fullerkati* and 90% *E. waccamawensis* test individuals (Table 2A). There was 100% amplification success for hemolymph extracts taken from both species from the Lake Waccamaw transects (Table 2B). In addition, sequences generated from amplified tissue and hemolymph of the same individual were aligned and compared to determine if there was any contamination in the hemolymph DNA, which is produced at very low yields and is prone to contamination. 16S rRNA gene tissue and hemolymph DNA sequences from the 5 *L. fullerkati*, 5 *E. waccamawensis*, and 4 *Leptodea ochracea* individuals compared were identical; therefore, there was no sign of contamination in the hemolymph DNA.

Table 2. Amplification success of hemolymph versus tissue DNA. Table 2A shows the amplification success of DNA extracted from both tissue and hemolymph samples of both *L. fullerkati* and *E. waccamawensis* individuals from the Lake Waccamaw enclosures. The numbers of each species have been adjusted based on the RFLP results; the two individuals that were morphologically misidentified as *L. fullerkati* were counted as *E. waccamawensis*. Table 2B shows the amplification success of hemolymph extracts from the Lake Waccamaw transects. Again, the species numbers were adjusted based on RFLP results. There were no DNA extractions taken from tissue of the transect animals.

A. Amplification Success of DNA from Lake Waccamaw Enclosures

Species	Amplification Success of Tissue Extracts	Amplification Success of Hemolymph Extracts
E. waccamawensis	29/29 (100%)	18/20 (90.0%)
L. fullerkati	19/19 (100%)	15/16 (93.8%)

B. Amplification Success of DNA from Lake Waccamaw Transects

Species	Amplification Success of Tissue Extracts	Amplification Success of Hemolymph Extracts
E. waccamawensis	n/a	31/31 (100%)
L. fullerkati	n/a	12/12 (100%)

Restriction Fragment Length Polymorphisms

Inconsistencies between species calls based on RFLPs and those based on morphology were due to three causes: morphological misidentifications of the species, and to ambiguities and polymorphisms at the restriction enzyme recognition sequences. Below, Figure 2 shows both consistent and inconsistent banding patterns.

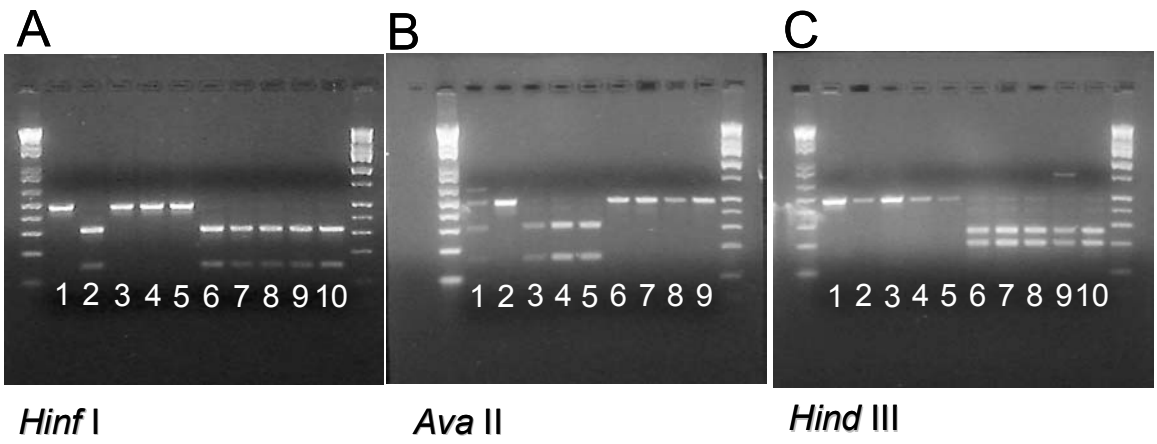

Figure 2. RFLP analysis of 16S rRNA gene PCR products. Gel A shows the products of a *Hinf* I digest. Lanes 1, 3-5 contain *L. fullerkati* amplicons; lanes 6-10 contain *E. waccamawensis* amplicons. The animal in lane 2 was morphologically misidentified as *L. fullerkati*. Gel B shows the products of an *Ava* II digest. Lanes 6-9 contain *E. waccamawensis* amplicons; lanes 1, 3-5 contain *L. fullerkati* amplicons. The animal in lane 2 was morphologically misidentified as *L. fullerkati*. Gel C shows the products of a *Hind* III digest. Lanes 1-5 contain *L. fullerkati* amplicons; lanes 6-10 contain *E. waccamawensis* amplicons. There are no misidentifications on Gel C.

A total of 69 *E. waccamawensis* individuals were typed using *Ava* II and *Hind* III; 39 were typed using *Hinf* I. For *L. fullerkati*, 43 individuals were typed using *Ava* II and *Hind* III, while 30 individuals were typed using *Hinf* I. Results from the RFLPs are shown in Table 3. *Ava* II produced banding patterns that were consistent with morphology for 98.6% of *E. waccamawensis* and for 93% of *L. fullerkati* individuals. *Hinf* I digests agreed with morphology for 92.3% of *E. waccamawensis* and for 90% *L. fullerkati*. Finally, *Hind* III produced consistent results for 97.1% of *E. waccamawensis* and for 90.7% of *L. fullerkati* (Table 3).

Table 3. RFLP results for Lake Waccamaw enclosures and transects. A. Shows the results for animals taken from Lake Waccamaw Enclosures from the mortality experiment. B. Shows the results for animals taken from the Lake Waccamaw transects. Ambiguous banding patterns appeared to be misidentifications; however, sequencing showed that these individuals were correctly identified. Polymorphisms at the recognition sequence of *Hinf* I was found by comparing aligned sequences from *E. waccamawensis* and *L. fullerkati*.

A. Lake Waccamaw Enclosures

Enzyme	Morphological species	Number typed	RFLP and morphology consistent	Number of morphological mis-ID's	Number of RFLP ambiguities	Number of RFLP site polymorphisms
***Ava* II**	*E. waccamawensis*	39	38	0	1	0
	L. fullerkati	30	28	2	0	0
***Hinf* I**	*E. waccamawensis*	39	36	0	0	3
	L.fullerkati	30	27	2	1	0
***Hind* III**	*E. waccamawensis*	39	37	0	2	0
	L. fullerkati	30	28	2	0	0

B. Lake Waccamaw Transects

Enzyme	Morphological species	Number typed	RFLP and morphology consistent	Number of morphological mis-ID's	Number of RFLP ambiguities	Number of RFLP site polymorphisms
***Ava* II**	*E. waccamawensis*	30	30	0	0	0
	L. fullerkati	13	12	1	0	0
***Hinf* I**	*E. waccamawensis*	0	N/A	N/A	N/A	N/A
	L.fullerkati	0	N/A	N/A	N/A	N/A
***Hind* III**	*E. waccamawensis*	30	30	0	0	0
	L. fullerkati	13	11	1	1	0

Individuals that produced inconsistent RFLPs were sequenced for the 16S rRNA gene to determine if the inconsistencies were due to misidentifications or if they were the result of ambiguities or polymorphisms in the restriction sites. The condensed 16S rRNA gene sequencing alignment for several individuals (Figure 3) shows the pronounced level of genetic distance between these species. Unlike the RFLPs, DNA sequence level distribution between the two species is never ambiguous (Figure 3).

	18	51	55	65	69	70	76	90	**98**	**99**	**100**	**101**	**102**
L. fullerkati LWD9	C	G	T	G	A	G	C	G	**G**	**A**	**C**	**C**	**C**
L. fullerkati Orange43	.	.	.	.	.	.	.	.	**.**	**.**	**.**	**.**	**.**
L. fullerkati Orange 45	.	.	.	.	.	.	.	.	**.**	**.**	**.**	**.**	**.**
L. fullerkati Orange44	.	.	.	.	.	.	.	.	**.**	**.**	**.**	**.**	**.**
L. fullerkati Orange40	.	.	.	.	.	.	.	.	**.**	**.**	**.**	**.**	**.**
L. fullerkati Orange41	.	.	.	.	.	.	.	.	**.**	**.**	**.**	**.**	**.**
E. waccamawensis Aqua51	?	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**T**
E. waccamawensis Aqua65	?	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**T**
E. waccamawensis Aqua40	?	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**T**
E. waccamawensis Aqua43h	T	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**.**
L. fullerkati Orange63	T	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**.**
L. fullerkati LWD22	T	A	C	T	G	A	A	A	**.**	**.**	**.**	**T**	**.**
E. waccamawensis Aqua52	T	A	C	T	G	A	A	A	**.**	**.**	**.**	**T**	**.**
L. fullerkati Orange71	T	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**.**
E. waccamawensis Aqua48	T	A	C	T	G	.	A	A	**.**	**.**	**.**	**T**	**.**

	104	107	150	152	155	156	157	163	166	172	173	174	175
L. fullerkati LWD9	A	G	G	T	C	C	A	T	G	T	T	G	A
L. fullerkati Orange43	.	.	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange 45	.	.	.	.	.	.	.	.	.	.	.	.	G
L. fullerkati Orange44	.	.	.	.	.	.	.	.	.	.	.	.	G
L. fullerkati Orange40	.	.	.	.	.	.	.	.	.	.	.	.	G
L. fullerkati Orange41	.	.	.	.	.	.	.	.	.	.	.	.	.
E. waccamawensis Aqua51	G	A	A	C	T	A	G	C	C	C	A	A	.
E. waccamawensis Aqua65	G	A	A	C	T	A	G	C	C	C	A	A	.
E. waccamawensis Aqua40	G	A	A	C	T	A	G	C	C	C	A	A	.
E. waccamawensis Aqua43h	G	A	A	C	T	A	G	C	C	C	A	A	.
L. fullerkati Orange63	G	A	A	C	T	A	G	C	C	C	A	A	.
L. fullerkati LWD22	G	A	A	C	T	A	G	C	C	C	A	A	.
E. waccamawensis Aqua52	G	A	A	C	T	A	G	C	C	C	A	A	.
L. fullerkati Orange71	G	A	A	C	T	A	G	C	C	C	A	A	G
E. waccamawensis Aqua48	G	A	A	C	T	A	G	C	C	C	A	A	G

	209	211	213	**214**	**215**	**216**	**217**	**218**	**219**	224	225	227	228
L. fullerkati LWD9	C	C	A	**A**	**A**	**A**	**C**	**T**	**T**	C	T	A	T
L. fullerkati Orange43	.	.	.	**.**	**.**	**.**	**.**	**.**	**.**	.	.	.	.
L. fullerkati Orange 45	.	.	.	**.**	**.**	**.**	**.**	**.**	**.**	.	.	.	.
L. fullerkati Orange44	.	.	.	**.**	**.**	**.**	**.**	**.**	**.**	.	.	.	.
L. fullerkati Orange40	.	.	.	**.**	**.**	**.**	**.**	**.**	**.**	.	.	.	.
L. fullerkati Orange41	.	.	.	**.**	**.**	**.**	**.**	**.**	**.**	.	.	.	.
E. waccamawensis Aqua51	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
E. waccamawensis Aqua65	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
E. waccamawensis Aqua40	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
E. waccamawensis Aqua43h	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
L. fullerkati Orange63	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
L. fullerkati LWD22	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
E. waccamawensis Aqua52	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
L. fullerkati Orange71	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C
E. waccamawensis Aqua48	A	A	C	**.**	**.**	**G**	**.**	**.**	**.**	T	C	C	C

	229	230	231	235	238	239	242	243	245	249	250	254
L. fullerkati LWD9	A	A	G	A	G	T	A	T	A	T	T	A
L. fullerkati Orange43	.	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange 45	.	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange44	.	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange40	.	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange41	.	.	.	.	.	.	.	.	.	.	.	.
E. waccamawensis Aqua51	C	T	T	T	A	C	C	C	T	A	C	.
E. waccamawensis Aqua65	C	T	T	T	A	C	C	C	T	A	C	.
E. waccamawensis Aqua40	C	T	T	T	A	C	C	C	T	A	C	.
E. waccamawensis Aqua43h	C	T	T	T	A	C	C	C	T	A	C	.
L. fullerkati Orange63	C	T	T	T	A	C	C	C	T	A	C	.
L. fullerkati LWD22	C	T	T	T	A	C	C	C	T	A	C	G
E. waccamawensis Aqua52	C	T	T	T	A	C	C	C	T	A	C	G
L. fullerkati Orange71	C	T	T	T	A	C	C	C	T	A	C	.
E. waccamawensis Aqua48	C	T	T	T	A	C	C	C	T	A	C	.

	256	**257**	**258**	**259**	**260**	297	300	301	351	360	382
L. fullerkati LWD9	**G**	**G**	**A**	**C**	**C**	A	T	G	T	A	C
L. fullerkati Orange43	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange 45	.	.	.	.	.	.	.	.	.	N	.
L. fullerkati Orange44	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange40	.	.	.	.	.	.	.	.	.	.	.
L. fullerkati Orange41	.	.	.	.	.	.	.	.	.	.	.
E. waccamawensis Aqua51	.	.	.	.	**A**	G	C	A	C	C	T
E. waccamawensis Aqua65	.	.	.	.	**A**	G	C	A	C	C	T
E. waccamawensis Aqua40	.	.	.	.	**A**	G	C	A	C	C	T
E. waccamawensis Aqua43h	.	.	.	.	**A**	G	C	A	?	?	?
L. fullerkati Orange63	.	.	.	.	**A**	G	C	A	C	C	T
L. fullerkati LWD22	.	.	.	.	**A**	G	C	A	C	C	T
E. waccamawensis Aqua52	.	.	.	.	**A**	G	C	A	C	C	T
L. fullerkati Orange71	.	.	.	.	**A**	G	C	A	C	T	T
E. waccamawensis Aqua48	.	.	.	.	**A**	G	C	A	C	T	T

Figure 3. Condensed alignment of 16S rRNA gene sequences. This alignment contains sequences from both *E. waccamawensis* and *L. fullerkati*. Shown on the alignment, in bold are the recognition sequences for the restriction enzymes *Hinf* I (98-102), *Hind* III (214-219), and *Ava* II (256-260). The individuals that were morphologically misidentified as *L. fullerkati* show bolded taxon names. Question marks represent missing data. Dots denote nucleotides identical to those for the first taxon at that position. Underlined column headings denote the nucleotide positions containing a variable site. Note the presence of fixed substitutions within the *Hind* III and *Ava* II sites that lead, respectively, to a gain and a loss of the restriction site in *L. fullerkati* individuals.

Ava II, *Hinf* I, and *Hind* III RFLPs detected two misidentifications; these *E. waccamawensis* individuals from the Lake Waccamaw enclosures were morphologically misidentified as *L. fullerkati* (Table 3). An additional misidentification was found among the Lake Waccamaw transect animals; *Ava* II and *Hind* III RFLPs detected an *E. waccamawensis* individual that was incorrectly identified morphologically as *L. fullerkati* (Table 3). DNA sequencing confirmed that these misidentifications were real, which is illustrated in the maximum likelihood phylogram (Figure 4). The sequences from *L. fullerkati* LWD22, Orange 63, and Orange 71 grouped within the clade containing the *E. waccamawensis* sequences and not with the clade containing the *L. fullerkati* sequences. *L. fullerkati* Orange 40 and *E. waccamawensis* Aqua 43, both of which produced consistent RFLP results, were included in the tree as reference sequences. The rest of the sequences were from individuals which produced inconsistent banding patterns, but that were found upon DNA sequence analysis to have been correctly identified based on morphology.

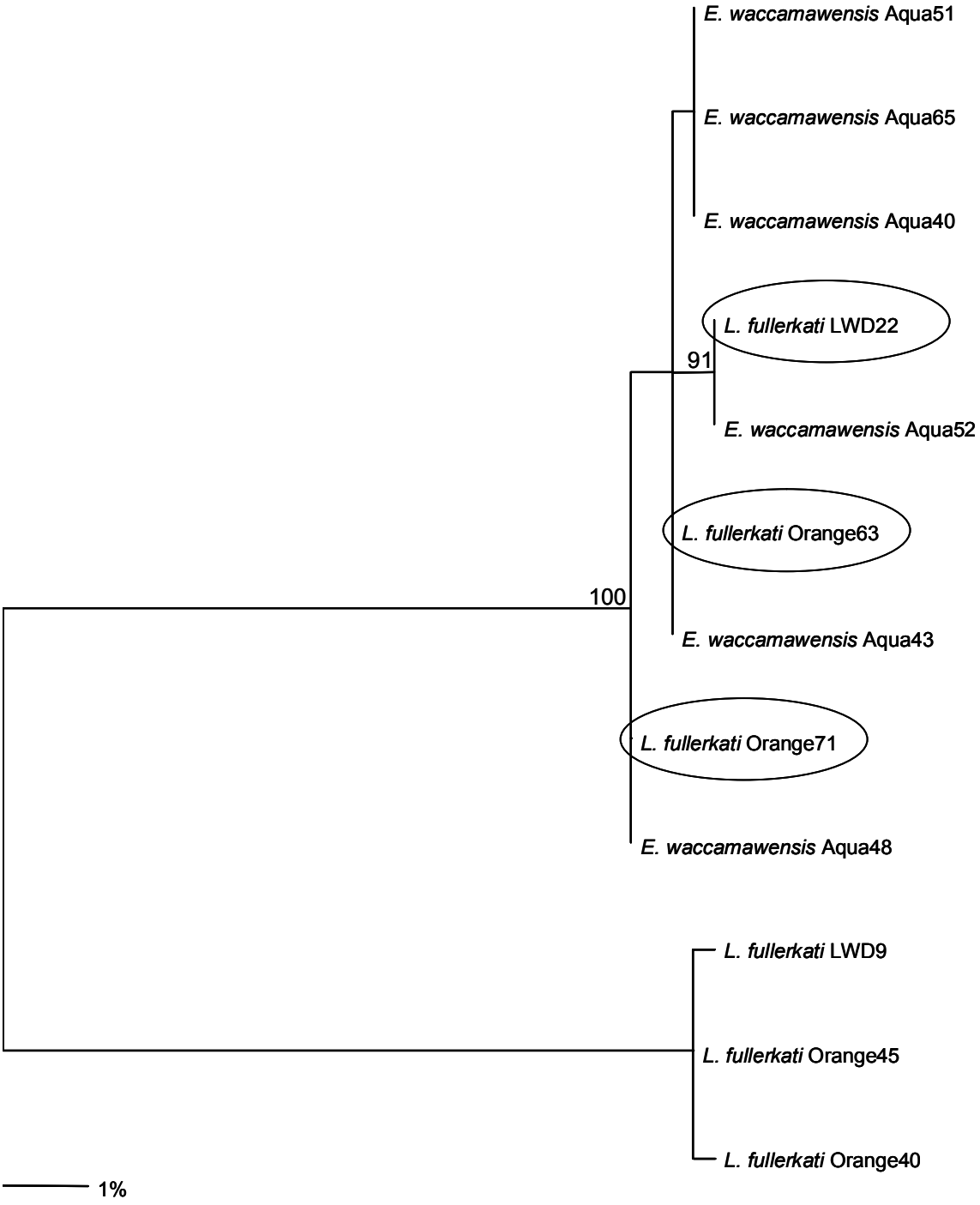

Figure 4. Maximum Likelihood phylogram of 16S rRNA gene sequences. The individuals included in this phylogram were those that produced inconsistent RFLP patterns, due to misidentification, ambiguities, or polymorphisms. *L. fullerkati* Orange 40 and *E. waccamawensis* Aqua 43 did produce consistent RFLP results; however, they were included in the tree to serve as reference sequences. Labels on nodes are bootstrap support values > 70. Circled individuals were misidentified morphologically as *L. fullerkati*.

Few ambiguous banding patterns were produced by *Ava* II (1), *Hinf* I (1), and *Hind* III (3). Unfortunately, three *E. waccamawensis* individuals contained polymorphisms at the *Hinf* I recognition sequence, which were identified in the condensed sequencing alignment as a secondary substitution at position 102 (Figure 3). At this position, *E. waccamawensis* Aqua51, Aqua65, and Aqua40 possessed a T instead of a C. Because of this substitution, *Hinf* I did not cut at the recognition sequence; therefore, the resulting banding patterns resembled that of *L. fullerkati* individuals. As seen in Figure 3, polymorphisms were detected at the recognition sequence of *E. waccamawensis* individuals, but not at the recognition sequence of *L. fullerkati* individuals. Also, neither species demonstrated polymorphic sites for *Ava* II or *Hind* III (Figure 3).

In addition, *Hinf* I, *Ava* II, and *Hind* III recognition sites were identified within *Elliptio*, *Lampsilis*, *Leptodea*, and *Uniomerus* DNA sequences from the combined gene alignment from Chapter 3. The recognition sequences of both *Hinf* I and *Hind* III were found in the majority of the *Elliptio* species (Table 4). The *Hind* III recognition sequence was also found in *Uniomerus* species. However, nearly all of the *Lampsilis* and *Leptodea* sequences contained the *Ava* II recognition sequence instead of those for *Hinf* I and *Hind* III (Table 4).

Table 4. Presence of *Hinf* I, *Ava* II, and *Hind* III recognition sites. Presented in this table are the species that share the recognition sequences of the above enzymes. Y denotes the presence of the enzyme recognition sequence and shaded cells denote the absence of the recognition sequence. "*E. complanata*" was identified morphologically as *E. complanata*, but is genetically indistinguishable from *U. carolinianus* (Chapter 3). * Pee Dee Lance.

Species	*Hinf* I Recognition Site	*Ava* II Recognition Site	*Hind* III Recognition Site
E. waccamawensis	Y		Y
E. congaraea	Y		Y
E. icterina	Y		Y
E. complanata	Y		Y
E. folliculata	Y		Y
E. producta	Y		Y
Elliptio sp. (PDL)*	Y		Y
E. fisheriana	Y		
E. crassidens	Y		Y
Fusconaia flava	Y		Y
Pleurobema clava	Y		Y
L. fullerkati		Y	
L. radiata		Y	
L. ovata		Y	
Leptodea ochracea	Y	Y	
Leptodea leptodon			
Uniomerus carolinianus			Y
" *E. complanata*"			Y
Uniomerus declivus			Y

DISCUSSION

On the whole, the NCWRC and NCDOT field biologists were able to successfully identify the animals collected from Lake Waccamaw based on morphology. RFLP and DNA sequencing analysis confirmed that nearly all of the individuals (97%) were correctly identified. However, RFLP results for *Hinf* I, *Hind* III, and *Ava* II demonstrated that 3 out of 112 individuals were misidentified based on morphology, which was confirmed by DNA sequencing. Overall, *Ava* II, *Hinf* I, and *Hind* III produced few ambiguous results. However, out of the three restriction enzymes, only *Ava* II and *Hind* III were found to be diagnostic for *E. waccamawensis* and *L. fullerkati*. Because of the sequence polymorphism at the recognition site of *Hinf* I, it is not diagnostic for all individuals of *E. waccamawensis* and cannot be used to differentiate between these two species. In addition, *Hinf* I, *Ava* II, and *Hind* III could also serve as diagnostic enzymes for other *Elliptio*, *Lampsilis*, *Leptodea*, and *Uniomerus* species (Table 4). In the future these enzymes could be employed in a PCR-RFLP key, similar to the one developed by White et al. (1994, 1996), to distinguish among North Carolina freshwater mussel species.

Identification of Lake Waccamaw endemics *Elliptio waccamawensis* and *Lampsilis fullerkati* is complicated by their convergence in shell morphology. Male *L. fullerkati* individuals can be distinguished from *E. waccamawensis* by slight differences, such as the sharpness of the posterior ridge (Bogan 2002). *E. waccamawensis* became listed as State Endangered species in July of 2002 (Bogan 2002). *L. fullerkati* is not yet on the endangered species list; however, it is considered to be a State Threatened species (Bogan 2002). In order to perform management of these species it is essential that they be correctly identified. For example, improper identification could result in inaccurate assessments of population health. The results of the present study demonstrated that RFLPs can be used by field biologists as a

secondary method of identification. A PCR-RFLP identification key, similar to that of White et al. (1996) can be designed for distinguishing between morphologically convergent freshwater mussels to supplement the morphologic identification keys that already exist. A panel of diagnostic enzymes can be developed, and the resulting banding patterns can be used the same way that morphologic characters are used in identification keys (White et al. 1996). Even though RFLPs will greatly aid in distinguishing between morphologically convergent species, morphology will still be necessary for initial identification. The RFLPs simply serve as a check to confirm the morphology call.

Further, the present study demonstrated that non-lethal methods of hemolymph extraction serve to produce DNA sequences identical to those from tissue samples, which agrees with the results of Henley et al. (2006). The results from the present study contribute to those of Gustafson et al. (2005) and Raley et al. (2006) by demonstrating that non-lethal hemolymph sampling could be done in the field and applied to downstream genetic analyses, and that coupled to the earlier enclosure results (McCartney and Wilbur 2007), these techniques are tolerated by the animals with no mortality under natural conditions. Overall, this thesis project presents new information that can be employed by field biologists in the census and conservation of freshwater mussels. The use of hemolymph sampling and PCR-RFLP will allow biologists to quickly confirm problematic morphological identifications, without having to sacrifice any of the animals they are trying to protect.

CHAPTER THREE: Phylogeny of unionid genera *Lampsilis* and *Elliptio* in southeastern North Carolina: determining species status of the Lake Waccamaw endemics, *Lampsilis fullerkati* and *Elliptio waccamawensis*

INTRODUCTION

Many studies of freshwater mussels have lead to a better understanding of their phylogenetics and systematics, but there remains much to uncover. Hoeh et al. (1998a) employed *cox1* to investigate the higher level taxonomy of bivalves. They determined that within the subclass Palaeoheterodonta, the order Unionoida is monophyletic and sister to the monophyletic order Trigonioida, which contains the only surviving genus, *Neotrigonia* (Hoeh et al. 1998a). Furthermore, Hoeh et al. (1998b, 2001) studied the higher level systematics within the order Unionoida using *cox1* and combined *cox1* and morphologic data, respectively. They determined that the order Unionoida contains the two superfamilies, Unionoidea and Etherioidea, which are further subdivided into families. The families Hyriidae, Margaritiferidae, Iridinidae, and Mycetopodidae each formed monophyletic groups, in addition to the superfamily Etherioidea. However, the superfamily Unionoidea and the family Unionidae were not monophyletic (Hoeh et al. 1998b; 2001). In addition, they determined that within the Unionoidea, the families Unionidae and Margaritiferidae formed sister taxa (Hoeh et al. 1998b; 2001).

The majority of freshwater mussels found in North America are from the family Unionidae (Graf and Cummings 2006); therefore, the rest of this paper will be focused on this family. Ambleminae is the dominant and most speciose subfamily of Unionidae in North America (Campbell et al. 2005). Campbell et al. (2005) used DNA sequences from mitochondrial genes (16S rRNA gene, *cox1*, and NADH dehydrogenase subunit 1 (*nad1*)) to study this subfamily. Their analysis contained 37 genera, and 126 species, of which 30 species were represented by the type specimen from which the genus is based (Campbell et al. 2005). Their results showed that most of the genera within the subfamily are polyphyletic. However,

their data did support the monophyly of the tribes Quadrulini, Lampsilini, and Pleurobemini. The tribe Amblemini could not be resolved from their analyses as a monophyletic assemblage (Campbell et al. 2005).

The tribes Pleurobemini and Lampsilini contain the genera *Elliptio* and *Lampsilis*, respectively (Campbell et al. 2005). The systematics of *Lampsilis* has been evaluated in the following studies: Roe et al. (2001), Kat (1983a), and Stiven and Alderman (1992). *Elliptio* was studied by Davis et al. (1981), Davis and Fuller (1981) and Davis (1984).

Roe et al. (2001) were interested in a special group of *Lampsilis* species, which produce a superconglutinate that contains many glochidia and is used to attract host fish. Their study demonstrated that these superconglutinate producing *Lampsilis* species formed a distinct monophyletic group within the *Lampsilis* genus (Roe et al. 2001). Roe and Hartfield (2005) later re-described this group as a new genus, *Hamiota*. Kat (1983a) used allozymes to study the systematics of populations of *Lampsilis radiata*, *Lampsilis splendida*, and the Lake Waccamaw endemic *Lampsilis* sp., which was later described by Johnson (1984) as *L. fullerkati*. Kat's results showed that *L. splendida* was related to both *L. radiata* and *Lampsilis* sp. (Kat 1983a). He also determined that *L. radiata* and *Lampsilis* sp. were genetically very similar; however, he concluded that they were distinct species and that their genetic similarity was due to the fact that they were recently separated and have not had time to significantly diverge (Kat 1983a). Stiven and Alderman (1992) studied *Lampsilis* species from North Carolina. Their main objective was to determine if the subspecies of *Lampsilis radiata* were genetically distinct and could be considered separate species. Their results demonstrated that the subspecies *L. radiata radiata* and *L. radiata conspicua* were not genetically distinct. In addition they found that these two subspecies were not significantly differentiated from *L. fullerkati*, the Lake Waccamaw endemic.

Contrary to Kat (1983a), Stiven and Alderman (1992) concluded that these species and subspecies should be considered together as populations of the single species *L. radiata*.

Davis et al. (1981), Davis and Fuller (1981), and Davis (1984) studied the systematics of *Elliptio* in North America. Davis (1984) constructed a network of the genetic distances from Davis et al. (1981) and Davis and Fuller (1981), in addition to new data. This network showed that there were three lineages of lanceolate *Elliptio* and three lineages of non-lanceolate *Elliptio*. The lanceolate lineages were as follows: 1. *E. producta* and *E. fisheriana*; 2. *E. lanceolata*; and 3. *E. folliculata* and *E. shepardiana* (Davis 1984). The non-lanceolate *Elliptio* lineages were 1. *E. congaraea*, 2. *E. complanata*, and 3. *E. mcmichaeli*. Lineage one was also related to *E. waccamawensis* from North Carolina, *E. mcmichaeli* and *E. arctata* from the Florida panhandle, and *E. complanata* from the Delmarva Peninsula (Davis 1984). The second lineage contained *E. complanata* individuals restricted to the east coast and St. Lawrence basin and *E. buckleyi* individuals from Florida. Lineage three contained *E. crassidens* from the Florida panhandle in addition to *E. mcmichaeli* individuals (Davis 1984). In addition to these relationships, Davis et al. (1981) determined that the Lake Waccamaw endemic species *E. waccamawensis* was genetically similar to the *Elliptio* found in local drainages. Further, Davis (1984) found evidence that *E. waccamawensis* was most similar to *E. complanata*.

The purpose of the present study was to utilize DNA extracted mostly from non-lethal hemolymph sampling and from limited tissue sampling from sacrificed specimens to construct a phylogeny of individuals of the genera *Lampsilis* and *Elliptio* from North Carolina. The main focus of the analysis was to determine the species status of the Lake Waccamaw endemics, *Lampsilis fullerkati* and *Elliptio waccamawensis* and to evaluate these results in comparison to those described above. In addition, by sampling freshwater mussels from the surrounding

drainages, we hope to determine to which species the Lake Waccamaw endemics are most closely related.

MATERIALS AND METHODS

Sample collection

Individuals from the genera *Lampsilis* and *Elliptio* were collected from several geographic locations in North Carolina, including Lake Waccamaw (LW), and the Waccamaw (WR), Yadkin/Pee Dee (YPD), Little Pee Dee (LPD), and Lumber (LR) Rivers (Figure 5). The number of individuals sampled for each location depended on water level conditions and their abundance, but generally, at least 5 specimens of each presumptive species from each site was the target sample size. With the aid of freshwater biologists from the NCWRC and the NCDOT, samples were collected and sorted by species. Once the animals were sorted, 20-100 µl of hemolymph was extracted from the adductor muscle using a 1 ml sterile syringe with a 27G1/2 needle. Shell length was measured with a caliper, and the animals were photographed and returned to the location from where they were obtained. A small number of individuals were sacrificed from each recognized operational taxonomic unit (OTU). These animals will be deposited and vouchered at the North Carolina Museum of Natural Sciences in Raleigh, NC in June 2007. The DNA sequences from these sacrificed animals will be deposited on Genbank as the present results are submitted for publication.

In addition, tissue samples from *Lampsilis radiata radiata* individuals from Big Creek, Lick Creek, Tar River, and Flat River that were previously collected in 2001 were included in the analyses. Big Creek is part of the Waccamaw River and Lick Creek is part of the Pee Dee River, both of which are in the major Yadkin/Pee Dee drainage system (Figure 5). The Tar River is part of the Tar-Pamlico drainage and the Flat River is part of the Neuse drainage.

Topotype specimens from Dr. Arthur E. Bogan at the North Carolina Museum of Natural Sciences were also included in this study. The species and their type localities are shown in Table 5, below.

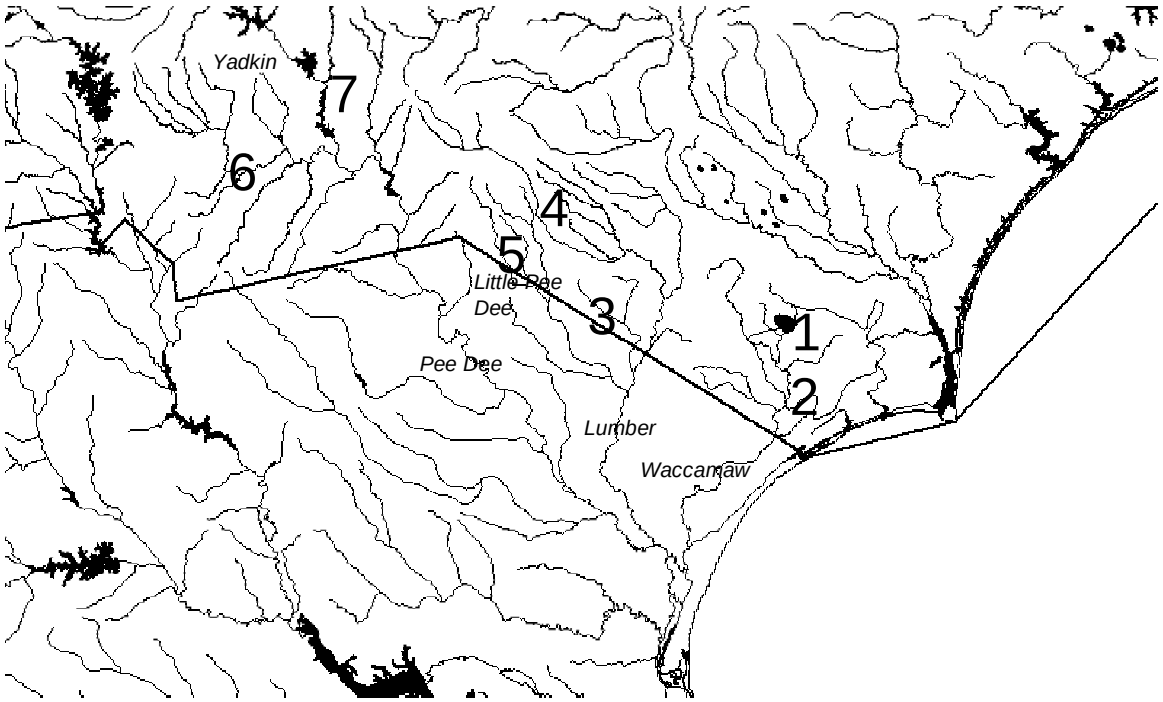

Figure 5. Freshwater mussel sampling locations from various watersheds in southeastern North Carolina. Sites in the Lumber River Basin are: (1) Lake Waccamaw, Columbus County; (2) Waccamaw River, Columbus Co; (3) Hog Swamp, Lumber River, Robeson Co.; (4) Richland Swamp and Hwy 71 Bridge, Lumber River, Robeson and Scotland Cos.; (5) Shoe Heel Creek, Little Pee Dee River, Robeson Co.; and in the Yadkin-Pee Dee River Basin: (6) Lick Creek, Pee Dee River, Anson Co.; (7) Morrow Mnt. State Park, Yadkin-Pee Dee River, Stanly Co. Major rivers are labeled in italics. Map of river drainages was obtained from the National Atlas of the U.S., Dept. of Interior.

Table 5. Topotype specimens and locations. This table includes a list of the topotype specimens obtained from Dr. Bogan at the North Carolina Museum of Natural Sciences. The table also includes the type localities from which these animals were originally described.

Species	NCSM Catalog #	Location	Latitude/Longitude
E. fisheriana	27696	Chesapeake Bay Basin/Chester River (Queen Annes Co.)	39.21904ºN, 75.90002ºW
E. fisheriana	27698	Chesapeake Bay Basin/Chester River (Queen Annes Co.)	39.2315ºN, 75.8563ºW
E. fisheriana	27699	Chesapeake Bay Basin/Chester River (Queen Annes Co.)	39.21904ºN, 75.90002ºW
E. fisheriana	41149	Savannah Basin/Savannah River (Screven Co.)	32.96421ºN, 81.51114ºW
E. icterina	41155	Savannah Basin/Savannah River (Screven Co.)	32.93857ºN, 81.50342ºW
E. icterina	41116	Savannah Basin/Savannah River (Aiken Co.)	33.45198ºN, 81.9205ºW
E. icterina	41130	Savannah Basin/Savannah River (Richmond Co.)	33.48906ºN, 81.98988ºW
E. producta	41135	Savannah Basin/Savannah River (Screven Co.)	32.89481ºN, 81.46989ºW
E. producta	26884	Savannah Basin/Savannah River (Allendale Co.)	0ºN, 0ºW
E. complanata	26965	Chesapeake Bay Basin/Potomac River (Montgomery Co.)	39.0004ºN, 77.2482ºW
E. complanata	26966	Chesapeake Bay Basin/Potomac River (Montgomery Co.)	39.0004ºN, 77.2482ºW

DNA extraction and purification

DNA was extracted from hemolymph using a PureGene DNA extraction kit (Gentra Systems, Minneapolis MN). The kit procedure and reagents were modified to accommodate small tissue volumes. A total of 200 µl of cell lysis solution and 1.5 µl of Proteinase K solution were combined with 50 µl of hemolymph. The samples were incubated overnight at 55ºC. Proteins, RNAs, and other cellular materials were separated from the DNA using 70 µl of protein precipitation solution and centrifugation. The DNA was then precipitated with 100% isopropanol and washed 70% ethanol. Once the ethanol was dried off, the DNA was resuspended in 35 µl sterile water. A StrataPrep PCR Purification Kit (Stratagene, La Jolla CA) was used to purify the DNA extracts through centrifugation. The DNA was bound to a column, washed, and eluted with 50 µl of sterile water. After elution, the samples were dried down and re-suspended in 10-15 µl of sterile water to obtain greater amplification success.

DNA was extracted from topotype tissue and from the adductor muscle tissue of the sacrificed animals using the same procedure as above, except that 200 µl of cell lysis solution and 1 µl of Proteinase K solution were combined with a 2-8 mg piece of tissue at the start of the procedure. Tissue extracts were also purified over StrataPrep columns; however, the cleaned tissue extracts did not need to be dried down and reconcentrated.

PCR amplification and purification

Three gene regions from the mitochondrial genome were amplified (cytochrome c oxidase subunit I gene (*cox1*), 16S ribosomal RNA gene (rRNA), and NADH dehydrogenase subunit 1 gene (*nad1*)). The reactions were carried out in 25 µl volumes, containing 1X PCR buffer with 1.5 mM $MgCl_2$ (ABI), 0.2 mM each dNTP, 0.5 µM each primer, and 1 U *AmpliTaq* polymerase (ABI). The primers and cycling conditions used to amplify these regions are shown

below (Table 6). Degenerate primers were created by modifying the Folmer et al. (1994) primers based on changes contained in *Lampsilis ornata* and *Mytilis galloprovincialis* sequences. In addition, for the 16S rRNA gene, a mitochondrial DNA touchdown program was used to amplify DNA extracts that were previously unsuccessful. For all three genes, *Taq*Gold (ABI) was used in place of traditional *AmpliTaq* to improve amplification success of failed or weak amplification products.

Resulting PCR products were purified using the StrataPrep PCR Purification Kit, as above. It was unnecessary to reconcentrate the purified PCR products before DNA sequencing.

Table 6. PCR primers, conditions, and product sizes. For *nad1*, the program used was from Campbell et al. (2005). The annealing temperature for the first five cycles was 45ºC, after which it increased to 50ºC for the next 25 cycles.

Gene Region	Forward Primer	Reverse Primer	Annealing Temp	Product Size
16S rRNA gene	16SUN693F (AGATAATGCCTGCCCAGTG)	16SUN1178R (CGGTCTTAACTCAGCTCGTGTA)	50ºC	~470bp
cox1	LCO1490 (Folmer et al. 1994) (GGTCAACAAATCATAAAGATATTGG) LCO1490D (GNTCNACNAATCATAARGATATTGG)	HCO2198 (Folmer et al. 1994) (TAAACTTCAGGGTGACCAAAAAATCA) HCO2198D (TAAACYTCAGGRTGNCCAAAAAATCA)	45ºC	~620bp
nad1	ND1_F (Serb et al. 2003) (TGGCAGAAAAGTGCATCAGATTTAAGC)	ND1_R (Serb et al. 2003) (GCTATTAGTAGGTCGTATCG)	45/50ºC	~770bp

DNA sequencing and alignment

Purified PCR products were sequenced in forward and reverse directions using Big Dye version 3.1 kits (ABI) and the above PCR primers at a 0.33μM concentration for the 16S rRNA gene, *cox1* (Folmer et al. 1994), and *nad1* (Serb et al. 2003) primers and a 1 μM concentration for the *cox1* degenerate primers. Sequencing reactions were loaded onto ABI 3100 and 3130 *xl* Genetic Analyzers. Sequences were edited using *Sequencher* (Gene Codes Corporation, Ann Arbor MI) and organized in *MacClade 4.0* (Maddison and Maddison 2000). Selected sequences from reference topotype specimens for each gene were obtained from Genbank (Table 7). The sequences were aligned using *Clustal X* (Thompson et al. 1994). After the alignment was created, the file was imported back into *MacClade* for editing. An alignment containing all of the individuals sequenced was created for each gene region, in addition to a combined alignment of all three gene regions. A smaller alignment of the combined gene regions was created that contained only *Elliptio* sequences and the two Pleurobemini reference sequences, *Pleurobema clava* and *Fusconaia flava* (This will be further referred to as the *Elliptio* only dataset). Appendix one contains a list of the genes sequenced for each individual in the dataset.

Table 7. Selected reference sequences from Genbank. Reference sequences from Lampsilini, Pleurobemini, and *Uniomerus* were taken from Genbank for each gene region. UAUC (University of Alabama Unionid Collection) sequences were from Campbell et al. (2005). Topotype specimens are denoted with a superscript T.

Taxon	**Accession Number**	**Reference**
16S rRNA gene		
Elliptio crassidens[T]	AY655034	UAUC3150
Fusconaia flava 1[T]	AY238481	Krebs et al. (2003)
Lampsilis ovata[T]	AY655048	UAUC108
Leptodea leptodon[T]	AY655050	UAUC135
Pleurobema clava[T]	AY655060	UAUC1477
Uniomerus declivus	AY655081	UAUC3290
cox1		
Elliptio crassidens[T]	AY613820	UAUC1493
Fusconaia flava 1[T]	AF231733	Bogan & Hoeh (2000)
Lampsilis ovata[T]	AY613826	UAUC108
Leptodea leptodon[T]	AY655003	UAUC135
Pleurobema clava[T]	AY655013	UAUC1477
Uniomerus declivus	AY613846	UAUC3290
nad1		
Elliptio crassidens[T]	AY613788	UAUC3150
Fusconaia flava 1[T]	AY613793	UAUC2864
Lampsilis ovata[T]	AY613797	UAUC1681
Leptodea leptodon[T]	AY655105	UAUC135
Pleurobema clava[T]	AY613802	UAUC1477

For the alignments of the protein coding regions, *cox1* and *nad1*, the flatworm mitochondrial genetic code from *MacClade* and the longest open reading frame were used to compare the amino acid sequences of the DNA sequences generated here to those of *Lampsilis ornata* (Serb and Lydeard 2003: *cox1* (124111983) and *nad1* (116709742)) in order to check sequence and alignment quality. For each alignment, the redundant sequences were merged in *MacClade*, to eliminate identical haplotypes.

Phylogenetic analysis

Phylogenetic trees were created using Bayesian analysis in *MrBayes version 3.2.1* (Huelsenbeck and Ronquist 2001). A best-fitting model of sequence evolution was determined for each alignment using *MrModeltest 2.2* (Nylander 2004). *MrModeltest* is similar to the program *Modeltest 3.06* (Posada and Crandall 1998); however, it only searches the models that *MrBayes* recognizes (Nylander 2006). The selected model and parameters for each analysis are shown below (Table 8). Each *MrBayes* analysis was run for 10,000 generations to determine the number of generations that would accumulate in an overnight (16 hour) run (Table 8). The number of generations it took to reach stabilization of the likelihood values was also determined for each analysis. Trees were collected every 100^{th} generation, and a burnin value of 10% of the total number of trees sampled was applied to each analysis. The burnin values were greater than the number of trees collected prior to the stabilization of the likelihood values.

Table 8. Models and parameters for *MrBayes* analyses. Models used during analysis were those identified using Hierarchical Likelihood Ratio Tests (hLRTs) in *MrModeltest*. Burnin values were calculated as 10% of the total number of trees sampled. Trees were sampled at a frequency of every 100th generation. GTR=general time reversible (Lanave et al. 1984). HKY (Hasegawa et al. 1985).

Dataset	Model Selected	Model Parameters	Number of Generations	Burnin values (# of trees discarded)
16s	GTR+I+G	I=0.4649 α=0.3678	2,510,000	2,510
cox1	GTR+I+G	I=0.6188 α=1.7705	1,510,000	1,510
nad1	GTR+I+G	I=0.5505 α=3.1035	1,510,000	1,510
Combined Gene	GTR+I+G	I=0.5720 α=1.3146	1,010,000	1,010
Combined *Elliptio*	HKY+I+G	I=0.5238 α=0.6951	1,010,000	1,010

In addition to *MrBayes*, maximum likelihood and maximum parsimony analyses were performed on the *Elliptio* only dataset using *PAUP* 4.0b10* (Swofford 2002) and based on the best-fit model from *MrModeltest*. The maximum likelihood analysis was performed using a full heuristic search with 10 random sequence additions and TBR branch swapping. Bootstrap analyses were performed using fast-stepwise additions and 100 replicates. Because the dataset contained several nearly identical sequences, it was too complicated for parsimony analysis to run to completion; the analysis would have infinitely swapped branches. Therefore, a parsimony analysis with simple stepwise addition and TBR branch swapping was run on *PAUP* until the computer ran out of memory. A strict consensus tree was created from the trees collected during the partial analysis.

Tests for monophyly

The status of the Lake Waccamaw endemic species *L. fullerkati* and *E. waccamawensis* was tested under the criterion of monophyly. The presence/absence of monophyly was determined using the Shimodaira-Hasegawa (1999) topology tests in *PAUP*. In order to perform these tests, the *MrBayes* consensus tree file from the combined gene analysis was compared to a tree file created in *MacClade*, in which individuals of *L. fullerkati* and *E. waccamawensis* were each forced to form monophyletic clades using the minimum number of topological changes to the true tree. Topology tests were performed, one at a time, comparing the true tree to the one in which either *L. fullerkati* or *E. waccamawensis* was forced to be monophyletic, using the resampling method RELL (Kishino et al. 1990). The tree with the highest likelihood was selected as the best tree.

RESULTS

Sequencing analysis

16S rRNA gene

The aligned 16S rRNA gene sequences contained two regions, totaling 16 characters that could not be aligned with confidence because of insertion-deletion mutations, and therefore, these sites were excluded from the analyses. Campbell et al. (2005) found similar regions, which were also excluded from their analyses. A total of 108 sequences were generated for 16S rRNA gene. Including the reference sequences the resulting alignment contained 114 sequences and 444 base pairs. After the identical sequences were merged, the alignment included 63 unique haplotypes.

cox1

The partial *cox1* gene dataset included 102 sequences of approximately 600 base pairs. There were no internal gaps present within the *cox1* alignment. *cox1* was more difficult to amplify than the 16S rRNA gene and required increasing magnesium concentrations in order to improve amplification success for some individuals. The completed alignment contained 108 sequences, including reference sequences, and 576 base pairs. The *cox1* alignment contained far fewer identical sequences than 16S rRNA gene; after merging these sequences, there remained 99 unique haplotypes.

nad1

nad1 sequences were generated using the primers from Serb et al. (2003) (Table 6). A total of 88 partial *nad1* sequences of 770 base pairs were generated. There were no internal gaps within the *nad1* alignment. *nad1* amplified more easily than *cox1*; however, none of the *Uniomerus* or related individuals would amplify at *nad1*. The completed *nad1* alignment, including reference sequences, was composed of 93 sequences and 767 base pairs. *nad1*

contained more identical haplotypes than *cox1*, but fewer than the 16S rRNA gene. There were 72 unique haplotypes in the *nad1* alignment after merging the redundant taxa.

Combined gene dataset

Aligned sequences from 16S rRNA gene, *cox1*, and *nad1* were appended together for each individual. However, some individuals did not have sequences from every gene region (Appendix one); therefore, these gaps were labeled as missing data. For the combined alignment, the insertion-deletion regions of the 16S rRNA gene were also excluded. Including reference sequences, the combined alignment contained a total of 115 sequences and 1787 base pairs. Of the 1787 base pairs, 1208 were conserved and 593 were variable. Identical sequences were merged, reducing the number of unique haplotypes to 111.

Combined *Elliptio* only

Only sequences from *Elliptio*, excluding those that grouped with *Uniomerus* were included in this dataset for 16S rRNA gene, *cox1*, and *nad1* gene regions. The purpose of this abridged alignment was to allow comparison of the results from different tree building methods, focused on relationships within *Elliptio*, which showed the greatest sensitivity to gene region and analysis details. The reference sequences *Pleurobema clava*, *Fusconaia flava*, and *Elliptio crassidens* were also included in this dataset. The 16S rRNA gene portion of this alignment contained insertion-deletion mutations that could not be aligned with confidence, which resulted in the exclusion of 6 characters. The data set included a total of 70 sequences and 1813 characters. Only one pair of identical sequences was found which reduced the number of unique haplotypes to 69.

Phylogenetic analysis

16S rRNA gene

The model specified for the analysis was a general time reversible (GTR) model with proportion of invariable sites (I) = 0.4649 and a gamma function shape parameter (α) = 0.3678. The analysis was run for a total of 2,510,000 generations, from which 25,100 trees were sampled. Ten percent of these trees were excluded before the creation of the consensus tree in order to ensure that no trees sampled before the stabilization of likelihood values were included (Table 8). The *MrBayes* consensus tree (Figure 6) was rooted with *Uniomerus declivus*, which is within the Ambleminae but outside of the tribes Lampsilini and Pleurobemini (Campbell et al. 2005). There were seven important groups within the tree (A-H Figure 6). Because of the frequency of identical haplotypes and number of individuals found within them, the haplotypes (Table 9) were given numbers and displayed on the tree to facilitate viewing. Clade A contained individuals from several *Elliptio* species, *E. waccamawensis*, *E. icterina*, *E. congaraea*, *E. complanata*, and the *E. crassidens* reference sequence. There was no evidence for any relationships among species within this clade. The next major group was composed of the lance forms of *Elliptio*: *E. folliculata*, *E. producta*, and *Elliptio* sp. (PDL), which is not yet described but is commonly referred to as the Pee Dee Lance. These species were included in subgroups B and C and Haplotype 44, which also contained the topotype specimens *E. fisheriana* NCSM41149 and *E. producta* NCSM41135, both from the Savannah River (Table 5). Clade D contained several *Elliptio* species, in addition to the basal Pleurobemini reference sequences *Fusconaia flava* and *Pleurobema clava*. Clade E, which grouped with Clade D, formed a monophyletic clade containing only *E. fisheriana* sequences from the Waccamaw and Little Pee Dee Rivers, in addition to the *E. fisheriana* topotypes NCSM27696, NCSM27698, and NCSM27699 (Table 5). Clades F and G contained sequences from the genus *Lampsilis*. Clade F was the major *Lampsilis* clade, which contained sequences from *Lampsilis fullerkati* from Lake

Waccamaw in addition to *Lampsilis radiata* individuals from the Yadkin/Pee Dee River, Big Creek, and the Flat River. Clade G contained individuals of *L. radiata* from the Yadkin/Pee Dee River from the Tar River and Lick Creek. This clade was differentiated from the other *L. radiata* sequences with strong support (100%); therefore, this is most likely a cryptic species. Within the major clade containing both *Lampsilis* clades F and G were the Lampsilini reference sequences *Leptodea leptodon* and *Lampsilis ovata*. Haplotype 35 contained the *Leptodea ochracea* sequences from Lake Waccamaw. The last clade (H) on the tree was comprised of *Uniomerus carolinianus* individuals from the Little Pee Dee and Lumber Rivers. In addition, it also included "*Elliptio complanata*" sequences from the Yadkin/Pee Dee and Lumber Rivers. These *E. complanata* individuals were most likely morphologically misidentified in the field during sample collection. Within this clade there was evidence for geographic separation between *U. carolinianus* individuals from the Little Pee Dee River and those from the Lumber River (96% support). Topotype specimen *E. icterina* NCSM41116 (Table 5) fell basal to the *Uniomerus* clade (H) with 100% support. Sequences from adductor muscle tissue of sacrificed specimens are denoted with a V before the sequence number and are in bold font on the tree. For the 16S rRNA gene phylogeny, all of these sacrificed sequences grouped with hemolymph sequences of the same species and sampling locations.

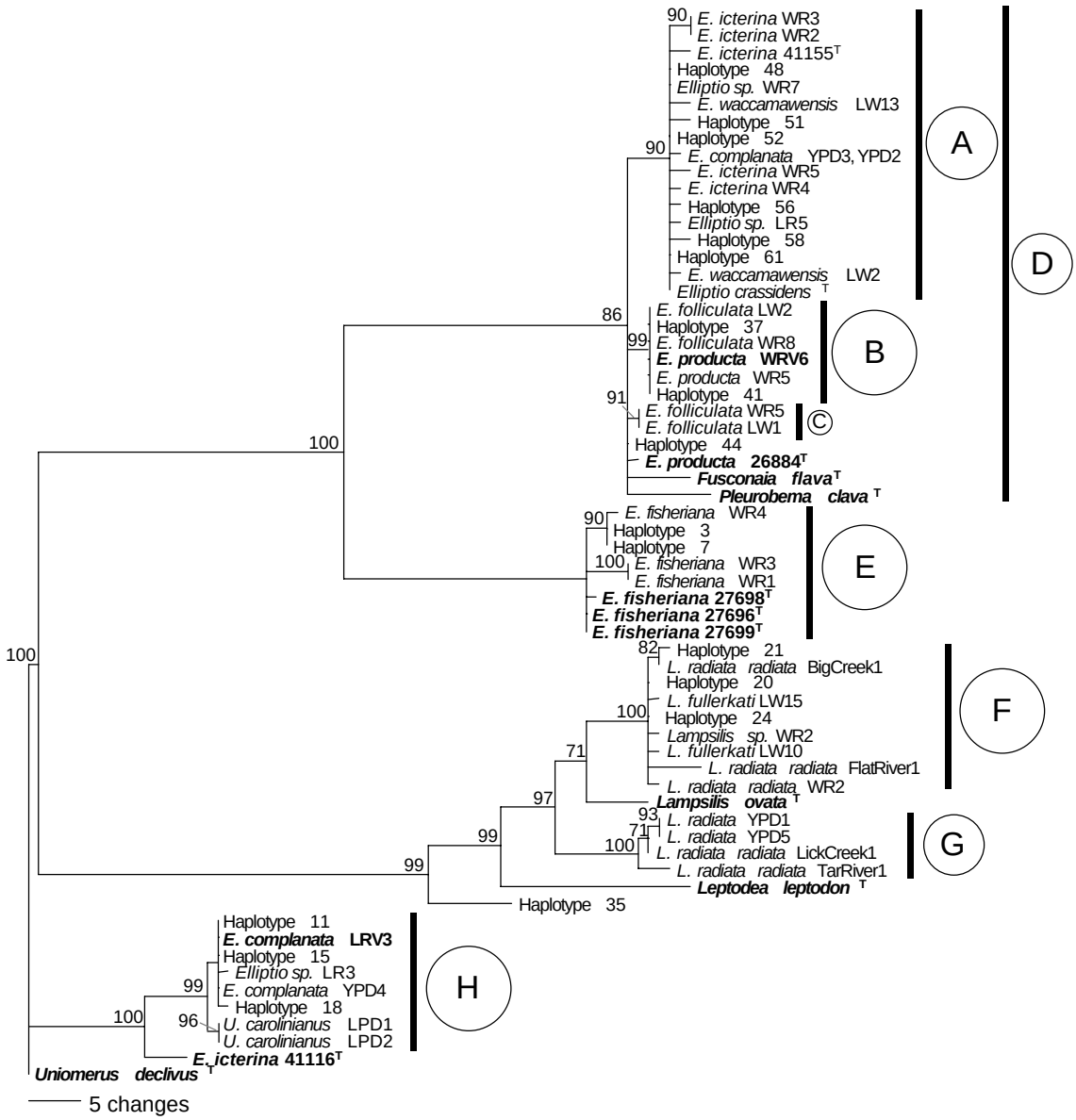

Figure 6. 16S rRNA gene *MrBayes* phylogeny. *MrBayes* clade probability values greater than 70% were included on the tree. A-H represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and adductor muscle tissue sequences from sacrificed animals are in bold. Topotypes are denoted with a superscript T. Sequences from adductor muscle tissue of sacrificed animals are denoted with a V preceding the sequence number. * Pee Dee Lance.

Table 9. 16S rRNA gene haplotypes. This table contains haplotypes, which were formed by merging redundant sequences. The sequences within each haplotype were identical. Sacrificed animal tissue and topotype sequences are in bold. Topotype specimens are denoted with a superscript T. Sequences from tissue of sacrificed animals are denoted with a V preceding the sequence number. * Pee Dee Lance.

Haplotype number	**Individuals Within Haplotype**
Haplotype 3	*E. fisheriana* LPD1, LPD2, LPD3, LPD4; *E. fisheriana* WR10, **WRV10**
Haplotype 7	***E. fisheriana* LPDV9**; *E. fisheriana* WR7
Haplotype 11	*E. complanata* YPD1; *Elliptio sp.* LR4
Haplotype 15	*E. complanata* LR3, LR4, **LRV2**, LR1 *Elliptio sp.* LR1, **LRV16**, LR2
Haplotype 18	*U. carolinianus* LR1, **LRV14**
Haplotype 20	*Lampsilis sp.* WR1; *L. radiata radiata* WR1; *L. fullerkati* LW12
Haplotype 21	*L. radiata* YPD2, YPD3
Haplotype 24	*L. fullerkati* LW11, LW3; *L. radiata* YPD4
Haplotype 35	*Leptodea ochracea* LW16, LW17, LW18
Haplotype 37	*Elliptio* sp. (PDL)* YPD3; *E. folliculata* WR3
Haplotype 41	*Elliptio sp.* YPD1; *Elliptio* sp. (PDL)* YPD1; *E. folliculata* WR4; ***E. producta* WRV5**
Haplotype 44	*E. folliculata* WR6**; *E. fisheriana* 41149^{T}**; ***E. producta* 41135^{T}**; *Elliptio* sp. (PDL)* YPD2
Haplotype 48	*E. waccamawensis* WR9, WR6, LW1, LW5, LW8; *Elliptio sp.* WR3, WR6; *E. congaraea* WR4, WR1, WR2, WR5; ***E. icterina* 41130^{T}**
Haplotype 51	***E. complanata* 26965^{T}; 26966^{T}**
Haplotype 52	*Lampsilis sp.* WR6, WR4, **WRV13**; *E. waccamawensis* WR10
Haplotype 56	*E. waccamawensis* WR5, WR1; *E. icterina* **WRV12**, WR1
Haplotype 58	*Lampsilis sp.* WR5; *Elliptio sp.* WR4
Haplotype 61	*Elliptio sp.* **LPDV17**, WR1; *E. congaraea* **WRV15**, WR3

cox1

The analysis was run with a model [GTR + (I=0.6188) + γ (α = 1.7705)] for a total of 1,510,000 generations, sampling a total of 15,100 trees. The first 1,510 trees were excluded from the consensus tree (Table 8). The phylogeny (Figure 7) was again rooted using the *Uniomerus declivus* sequence. Clade A contained individuals of *E. congaraea* from the Waccamaw River and *E. waccamawensis* from Lake Waccamaw and the Waccamaw River. This clade also contained *Lampsilis* sp. individuals from the Waccamaw River, which were most likely morphologically misidentified as *Lampsilis*. Individuals of *Elliptio complanata* from the Yadkin/Pee, and the topotype specimens NCSM26965 and NCSM26966 from the Potomac River in the Chesapeake Bay Basin (Table 5) were basal to clade A. Groups B and C contained sequences from *Elliptio icterina* and *E. waccamawensis* from the Waccamaw River. Clade D contained several *Elliptio* species, which were described by Davis et al. (1981) as the *E. complanata*-like forms of *Elliptio*. Clade E was composed of sequences from the lance forms of *Elliptio*: *E. producta*, *E. folliculata*, and *Elliptio* sp. (PDL). This clade also contained the topotype specimens *E. fisheriana* NCSM41149 and *E. producta* NCSM41135 (Table 5). Clade F was a monophyletic clade containing *Elliptio fisheriana* from the Waccamaw and Little Pee Dee Rivers, and also the topotype specimen of *E. fisheriana* NCSM27698 from the Chester River in the Chesapeake Bay Basin (Table 5). *Pleurobema clava* and *Fusconaia flava* reference sequences grouped together between clades E and F. The *Lampsilis* sequences were found in clades G and I. Clade G was the major *Lampsilis* clade and clade I contained the cryptic *Lampsilis* species. *Lampsilis ovata* fell outside of clade I. Clade H, which contained *Leptodea ochracea* sequences from Lake Waccamaw, fell between clades G and I, making *Lampsilis* paraphyletic. The *Leptodea leptodon* reference sequence was basal to the clade containing

clades G-I. Clade J was comprised of *Uniomerus carolinianus* sequences in addition to "*E. complanata*" sequences, as in the 16S rRNA gene tree. Unlike the 16S rRNA gene tree, there was no evidence for geographic differentiation between *U. carolinianus* from the Lumber River and those from the Little Pee Dee River. The topotype specimen *E. icterina* NCSM41116 from the Savannah River (Table 5) grouped with clade J. As seen in the 16S rRNA gene phylogeny, the sacrificed animal sequences from each clade grouped with hemolymph sequences of the same species and the same locations.

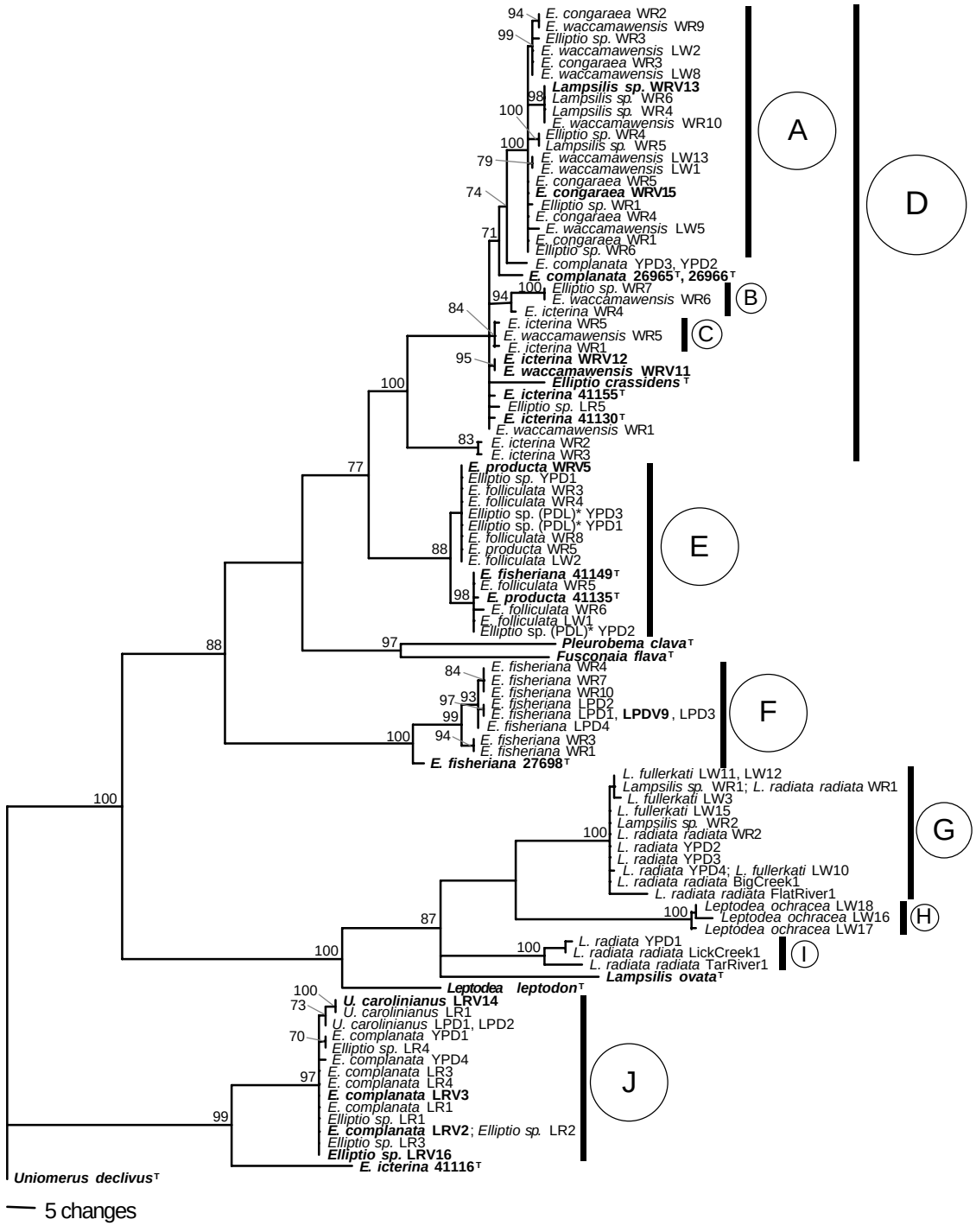

Figure 7. *cox1 MrBayes* phylogeny. *MrBayes* clade probability values greater than 70% were included on the tree. A-J represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and sequences from adductor muscle tissue of sacrificed animals are in bold. Topotype species are denoted with a superscript T. Sacrificed animal sequences are denoted with a V preceding the sequence number. * Pee Dee Lance.

nad1

The analysis was run with a model [GTR + (I=0.5505) + γ (α = 3.1035)] for 1,510,000 generations, from which 15,100 trees were sampled. 1,510 trees were discarded before the formation of the consensus tree (Table 8). This phylogeny (Figure 8) was midpoint rooted because no *Uniomerus* individuals were successfully amplified or sequenced for *nad1* with the primers used. Similar to the *cox1* phylogeny, clade A contained *E. congaraea* from the Waccamaw River and *E. waccamawensis* from Lake Waccamaw and the Waccamaw River. Misidentified *Lampsilis* sp. individuals were present in clade A as well. Yadkin/Pee Dee and topotype specimen *E. complanata* individuals were basal to clade A, similar to *cox1*. *E. icterina* and *E. waccamawensis* sequences from the Waccamaw River grouped together in clade B. Similar to the *cox1* tree, clade C contained the *Elliptio* species of the *E. complanata*-like form (Davis et al. 1981). The lance forms of *Elliptio* were contained in clade D. Similar to *cox1*, this group formed a single distinct clade with high support (100%), which was not seen in the 16S rRNA gene phylogeny. Also, *E. fisheriana* formed a monophyletic clade (E), containing individuals from the Waccamaw River, the Little Pee Dee River, and *E. fisheriana* topotype specimens (NCSM27696, NCSM27698, and NCSM27699) from the Chester River in the Chesapeake Bay Basin (Table 5). The reference sequences, *Pleurobema clava* and *Fusconaia flava* came out basal to the rest of the *Elliptio* clades. Clade F contained the *Leptodea ochracea* sequences from Lake Waccamaw, which grouped with the *Leptodea leptodon* reference sequence. The *Lampsilis* sequences were found in clades G and H. These clades were similar to the *Lampsilis* clades in both the 16S rRNA gene and *cox1* phylogenies. Clade G was the main *Lampsilis* clade, which contained *L. fullerkati* from Lake Waccamaw and *L. radiata* from the Yadkin/Pee Dee River, the Waccamaw River, Big Creek, and the Flat River. The cryptic

Lampsilis species sequences formed clade H with 100% support. Similar to the *cox1* phylogeny, the *Lampsilis ovata* reference sequence grouped with the cryptic *Lampsilis* species clade; however, this was not found in the 16S rRNA gene phylogeny. In addition to the 16S rRNA gene and *cox1* phylogenies, the sequences from tissue of the sacrificed animals from each clade in the *nad1* phylogeny grouped with the hemolymph sequences of the same species and sampling location.

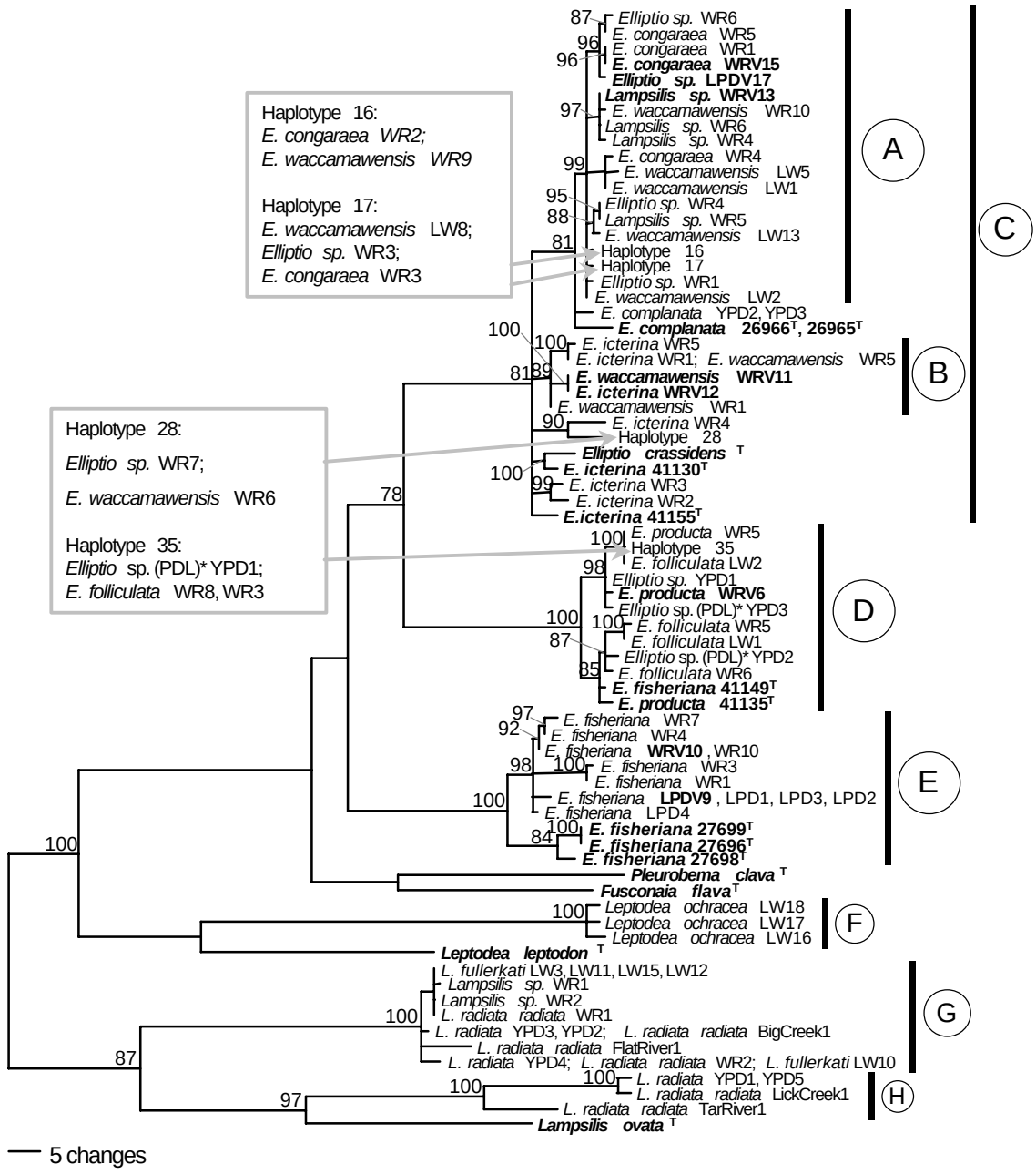

Figure 8. *nad1 MrBayes* phylogeny. *MrBayes* clade probability values greater than 70% were included on the tree. A-H represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and sacrificed animal tissue sequences are in bold. Topotypes are denoted with a superscript T. Sacrificed animal sequences are denoted with a V preceding the sequence number. * Pee Dee Lance.

Combined analysis

The combined analysis was run using a model [GTR + (I = 0.5720) + γ (α = 1.3146)] for a total of 1,010,000 generations. From these, a total of 10,100 trees were sampled. 1,010 trees were excluded from the consensus tree file (Table 8). Support values greater than 70 percent for the major clades were displayed at the nodes of the rectangular cladogram (Figure 9). Similar to *cox1* and *nad1* phylogenies, there was a clade composed of *E. congaraea* sequences from the Waccamaw River and *E. waccamawensis* sequences from Lake Waccamaw and the Waccamaw River (Clade A). *E. complanata* from the Yadkin/Pee Dee River and the *E. complanata* topotypes, NCSM26965 and NCSM26966 from the Potomac River (Table 5) grouped basal to clade A. Groups B and C contained *E. icterina* and *E. waccamawensis* sequences from the Waccamaw River. Unlike the *cox1* and *nad1* trees, there was not a defined clade containing all of the *E. complanata*-like *Elliptio* sequences; instead, these sequences fell out at the bottom of the tree. The lance forms of *Elliptio* all grouped in clade D, which contained two subclades. One of the subclades was supported (100%), the other was not supported. These clades contained sequences of *E. folliculata* from Lake Waccamaw and the Waccamaw River, *E. producta* from the Waccamaw River, and *Elliptio* sp. (PDL) from the Yadkin/Pee Dee River. The unsupported subclade also contained the topotype sequences *E. producta* NCSM41135, *E. producta* NCSM26884, and *E. fisheriana* NCSM41149 (Table 5). As in all other phylogenies, there was a monophyletic clade containing *E. fisheriana* sequences (Clade E). This clade also included the *E. fisheriana* topotype sequences NCSM27696, NCSM27698, and NCSM27699 from the Chester River (Table 5). *Fusconaia flava* and *Pleurobema clava* grouped together and outside of the *Elliptio* clades. Clades F and G were composed of *Lampsilis* sequences. Clade F was the major *Lampsilis* clade, which included *L. fullerkati* from Lake Waccamaw, and *L.*

radiata from the Yadkin/Pee Dee River, the Waccamaw River, Big Creek, and the Flat River. Clade G contained the cryptic *Lampsilis* species, with individuals from the Yadkin/Pee Dee River, Tar River, and Lick Creek. *Leptodea ochracea* sequences from Lake Waccamaw were found in clade H. The *Lampsilis ovata* reference sequence grouped with the cryptic *Lampsilis* species, and the *Leptodea leptodon* reference sequence grouped with the clade containing all *Lampsilis* species. *Uniomerus carolinianus* from the Little Pee Dee and Lumber Rivers grouped with misidentified "*E. complanata*" sequences from the Lumber and Yadkin/Pee Dee Rivers in clade I. Similar to the 16S rRNA gene and *cox1* phylogenies, *E. icterina* NCSM41116 from the Savannah River (Table 5) grouped outside of clade I. In addition, as was seen in each individual gene phylogeny, the sequences from tissue of sacrificed animals grouped with hemolymph sequences of the same species and same locations.

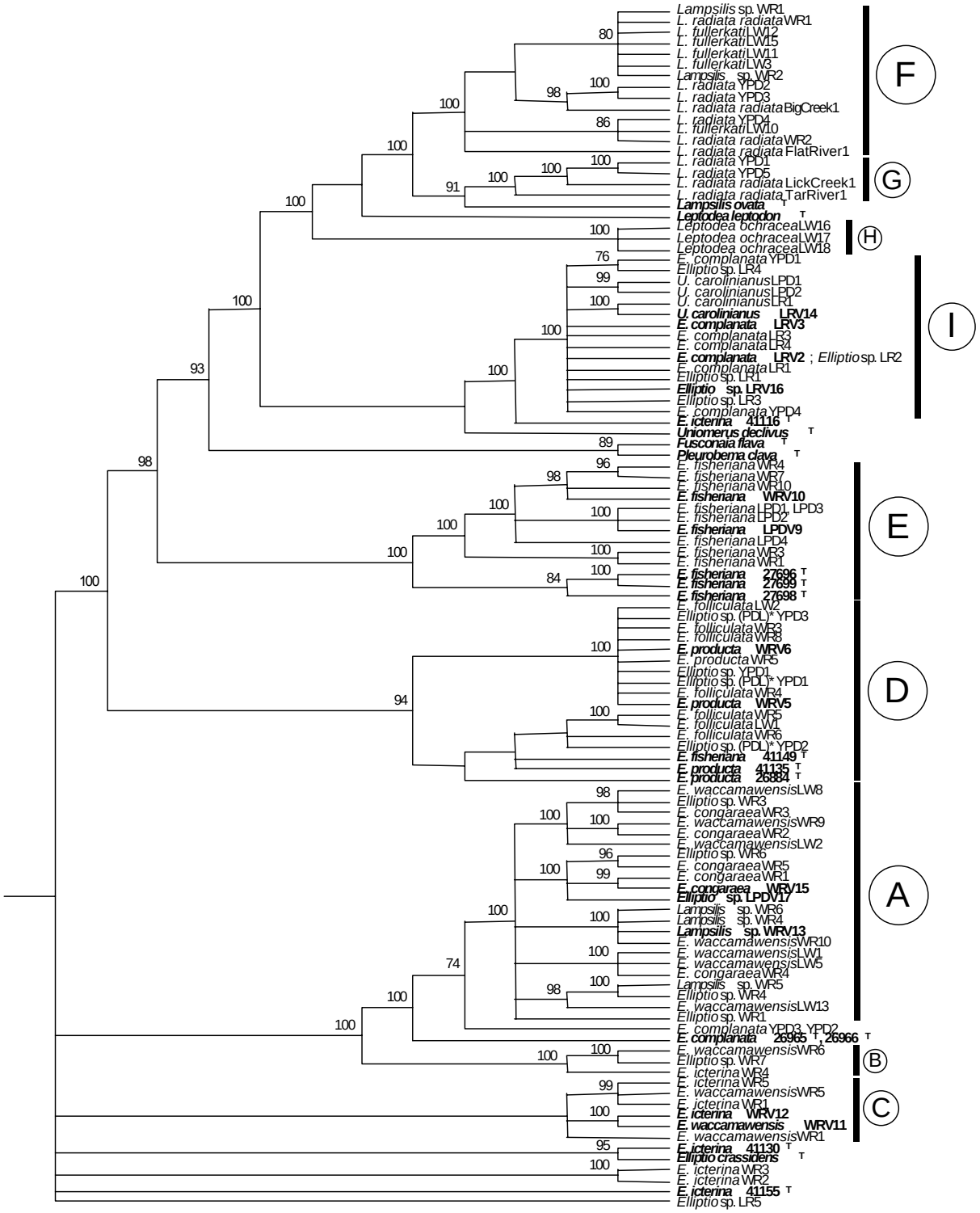

Figure 9. Combined *MrBayes* phylogeny of 16S rRNA gene, *cox1*, and *nad1* sequences. *MrBayes* clade probability values greater than 70% were included on the tree. A-I represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and sacrificed animal tissue sequences are in bold. Topotypes are denoted with a superscript T. Sacrificed animal sequences are denoted with a V preceding the sequence number. * Pee Dee Lance.

Combined *Elliptio* only

The analysis was run using a model [HKY + (I = 0.5238) + γ (α = 0.6951)] for a total of 1,010,000 generations, sampling 10,100 total trees. Of these trees, 1,010 were excluded from the formation of the consensus file (Table 8). The phylogeny (Figure 10) was midpoint rooted. Clade A contained sequences of *E. congaraea* from the Waccamaw River and *E. waccamawensis* from Lake Waccamaw and the Waccamaw River. *E. complanata* from the Yadkin/Pee Dee River and *E. complanata* topotype specimens NCSM26965 and NCSM26966 (Table 5) grouped basal to clade A. Groups B and C contained *E. waccamawensis* and *E. icterina* individuals from the Waccamaw River. Clade D was composed of the *E. complanata*-like *Elliptio* sequences (Davis et al. (1981). The lance forms of *Elliptio* were found in clade E, which contained two subclades. Clade E contained *E. producta* from the Waccamaw River, *E. folliculata* from Lake Waccamaw and the Waccamaw River, *Elliptio* sp. (PDL) from the Yadkin/Pee Dee River, and the topotype specimens *E. producta* NCSM41135, *E. producta* NCSM26884, and *E. fisheriana* NCSM41149 (Table 5). Once again, *E. fisheriana* formed a monophyletic clade (F). Within the *E. fisheriana* clade there was some suggestion of geographic variation between the *E. fisheriana* from the Waccamaw River and *E. fisheriana* from the Little Pee Dee River. However, *E. fisheriana* LPD4 did not fall in with the rest of the Little Pee Dee River *E. fisheriana* individuals. The *E. fisheriana* topotype specimens NCSM27696, NCSM 27698, and NCSM27699 (Table 5) grouped together (G) and with clade F. The *Pleurobema clava* and *Fusconaia flava* reference sequences grouped together at the bottom of the tree. The sacrificed animal tissue sequences grouped with hemolymph sequences of the same species and sampling locations.

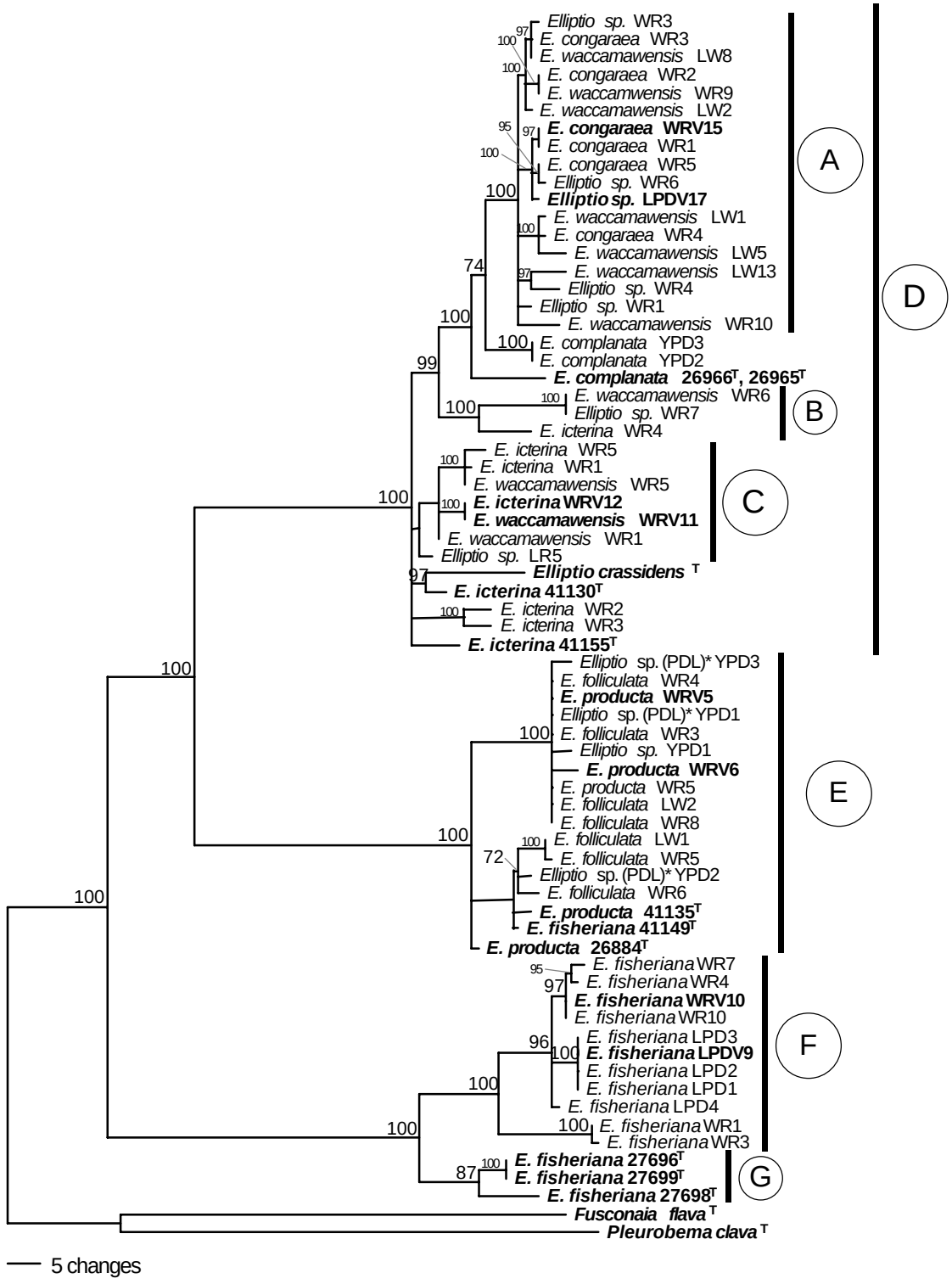

Figure 10. Combined *MrBayes* phylogeny of 16S rRNA gene, *cox1*, and *nad1 Elliptio* sequences. Support values greater than 70% were included on the tree. A-G represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and tissue sequences from sacrificed animals are in bold. Topotypes are denoted with a superscript T. Sacrificed animal sequences are denoted with a V preceding the sequence number. * Pee Dee Lance.

The maximum likelihood analysis of the *Elliptio* only dataset, which was run using the model above, 10 random sequence additions, and TBR branch swapping, produced nine equally probable trees. These trees were summarized in a strict consensus tree (Figure 11). In the consensus tree, clade A contained individuals of *E. congaraea* from the Waccamaw River and *E. waccamawensis* from Lake Waccamaw and the Waccamaw River. *E. complanata* individuals from the Yadkin/Pee Dee River and *E. complanata* topotypes NCSM26965 and NCSM26966 (Table 5) were basal to clade A. Groups B and C were composed of individuals from *E. icterina* and *E. waccamawensis*, both from the Waccamaw River. Overall, clade D contained the *Elliptio* sequences of the *E. complanata*-like form (Davis et al. 1981). *E. fisheriana* formed a monophyletic clade, which was composed of two subclades. One subclade contained the topotype specimens NCSM27696, NCSM27698, and NCSM27699 (E), and the other contained *E. fisheriana* sequences from the Waccamaw and Little Pee Dee Rivers (F). As seen in the *MrBayes* phylogeny above (Figure 10) there was some suggestion of geographic variation within clade F; all individuals from the Little Pee Dee River, except LPD4 grouped together with 98% support. Clades G and H contained the lance forms of *Elliptio*, in addition to two basal sequences, *E. producta* WRV6 and *Elliptio* sp. (PDL) YPD3. *Pleurobema clava* and *Fusconaia flava* reference sequences formed a clade which branched off the *E. fisheriana* clade. As seen in the *MrBayes* phylogeny, the sacrificed animal sequences grouped with hemolymph sequences of the same species and locations.

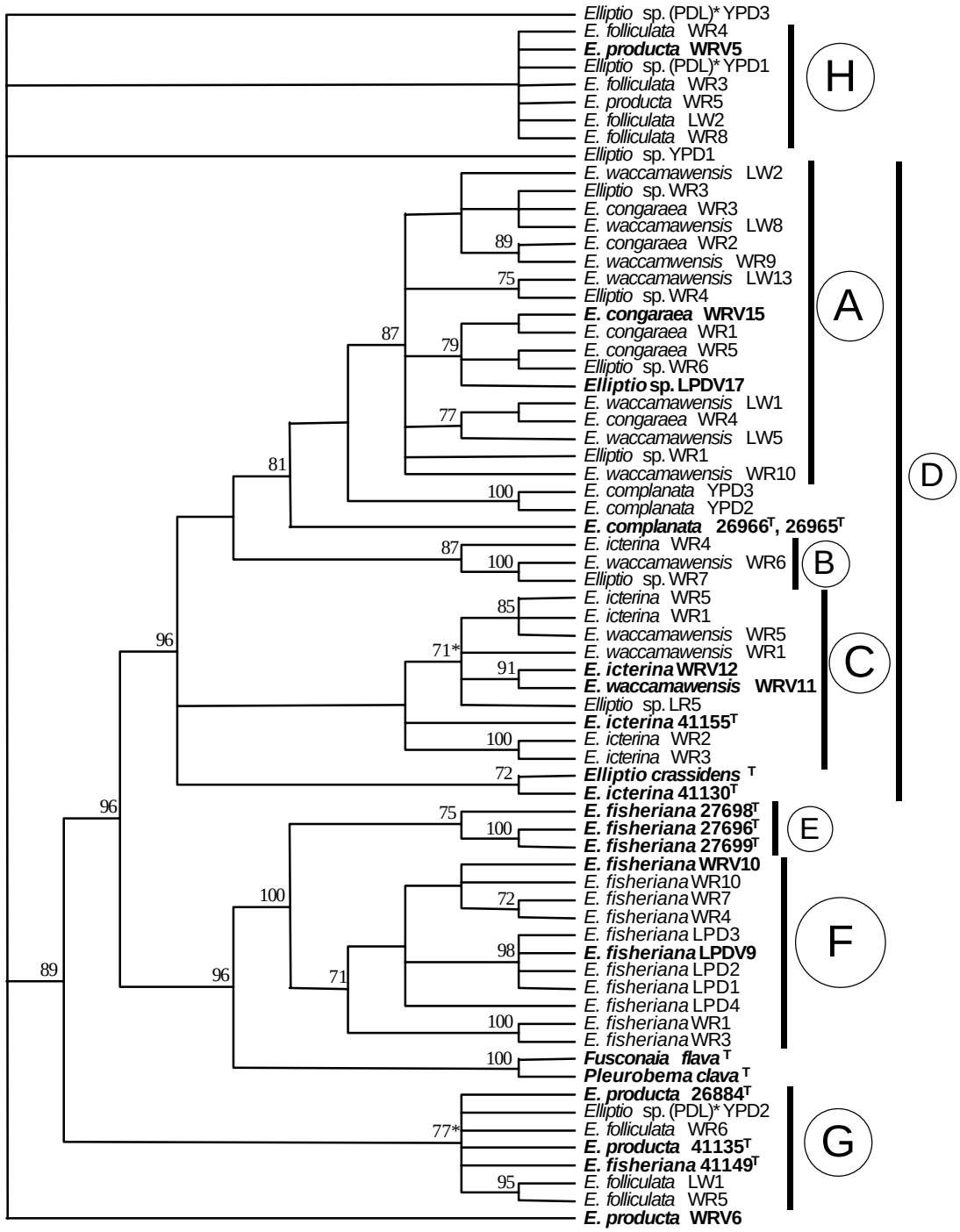

Figure 11. Combined maximum likelihood phylogeny of 16S rRNA gene, *cox1*, and *nad1 Elliptio* sequences. Strict consensus of 9 maximum likelihood trees produced during analysis. Bootstrap support values greater than 70% are included on the tree. Asterisked support values represent clades not supported in the *MrBayes* phylogeny. A-H represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and tissue sequences from sacrificed animals are in bold. Topotypes are denoted with a superscript T. Sacrificed animal sequences are denoted with a V preceding the sequence number. * Pee Dee Lance.

There were a couple of subtle differences in topology between the strict consensus maximum likelihood tree and the *MrBayes* tree. First, there were differences between groups formed by the lance forms of *Elliptio*. *Elliptio* sp. (PDL) YPD3 and *E. producta* WRV6 were basal sequences, instead of falling within the clade containing all the lance forms of *Elliptio* (Figure 10: Clade E; Figure 11: Clades G and H). As in the *MrBayes* tree, there were two subgroups of the lance forms; however, instead of grouping together in one clade, they formed two basal clades (Figure 11: Clades G and H). The second subgroup in the *MrBayes* phylogeny supported the grouping of *Elliptio* sp. (PDL) YPD2 and *E. folliculata* WR6 to the group that contained *E. folliculata* LW1 and *E. folliculata* WR5 with 72% (Figure 10). However, this relationship was not seen in the maximum likelihood consensus tree, which only supported the grouping of *E. folliculata* LW1 with *E. folliculata* WR5; the rest of the sequences came out basal within clade G (Figure 11).

The second difference between the phylogenies related to the placement of the *Fusconaia flava* and *Pleurobema clava* group. In the *MrBayes* phylogeny, this group fell out at the base of the tree (Figure 10). In the maximum likelihood phylogeny this clade grouped with the major *E. fisheriana* clade (Figure 11). Of course, in the combined analysis of all taxa (Figure 9), the placement of these species outside of and basal to *Elliptio* was highly supported.

For the most part, the maximum likelihood bootstrap support values agreed with the *MrBayes* clade probability values. There were only six cases out of 32 (19%) in which the *MrBayes* probability values were considered significant, while the maximum likelihood values were nonsignificant (less than 70% support). However, in all of these cases the maximum likelihood values were greater than 50%. The asterisked bootstrap values denoted clades that were supported by maximum likelihood analysis but not by the *MrBayes* analysis (Figure 10).

The *MrBayes* clade probabilities were more than 5% greater than the maximum likelihood bootstrap values in 20/32 cases (63%). For six of the shared clades, the *MrBayes* values were equal to the maximum likelihood values (19%). There were only two cases in which the maximum likelihood values were greater than the *MrBayes* values.

Maximum parsimony could not be completed with the combined *Elliptio* dataset because there were too many sequences that shared minor differences in addition to sequences that were distantly related. The analysis would never finish because for every branch swapping combination that was completed, the number of combinations remaining to be swapped increased. Therefore, a maximum parsimony analysis with simple stepwise addition and TBR branch swapping was run until the computer ran out of memory. At the end of the partial analysis, 447 trees were saved, and a strict consensus tree of those 447 trees was created (Figure 12). Bootstrap analyses could not be performed for the same reasons as above. Therefore, only the topologies of the maximum parsimony consensus tree and the *MrBayes* phylogeny could be compared.

In the consensus tree, clade A contained *E. congaraea* sequences from the Waccamaw River and *E. waccamawensis* sequences from Lake Waccamaw and the Waccamaw River. Both *E. complanata* from the Yadkin/Pee Dee River and the *E. complanata* topotypes NCSM26965 and NCSM26966 (Table 5) grouped basal to clade A. Groups B, C, and D contained *E. waccamawensis* and *E. icterina* individuals from the Waccamaw River. As in the other trees, the *Elliptio* sequences of the *E. complanata*-like form (Davis et al. 1981) fell within a single clade (E). Similar to the *MrBayes* and maximum likelihood analyses, *E. fisheriana* formed a monophyletic clade, which was composed of two subgroups (F and G). Clade F contained *E. fisheriana* sequences sampled from the Waccamaw and Little Pee Dee Rivers. As seen in the

MrBayes and maximum likelihood trees, the Little Pee Dee River sequences formed a distinct group, except for LPD4. Clade G contained the *E. fisheriana* topotype specimens NCSM27696, NCSM27698, and NCSM27699 (Table 5). The lance *Elliptio* sequences were found in groups H and I, which were not distinct clades, but instead were two groups of basal sequences. The only internal relationship shown among these sequences was the clade formed by *E. folliculata* LW1 and *E. folliculata* WR5. Similar to the maximum likelihood tree, *Pleurobema clava* and *Fusconaia flava* reference sequences grouped together and with the major *E. fisheriana* clade. Again, tissue sequences from the sacrificed animals grouped with hemolymph sequences of the same species and locations.

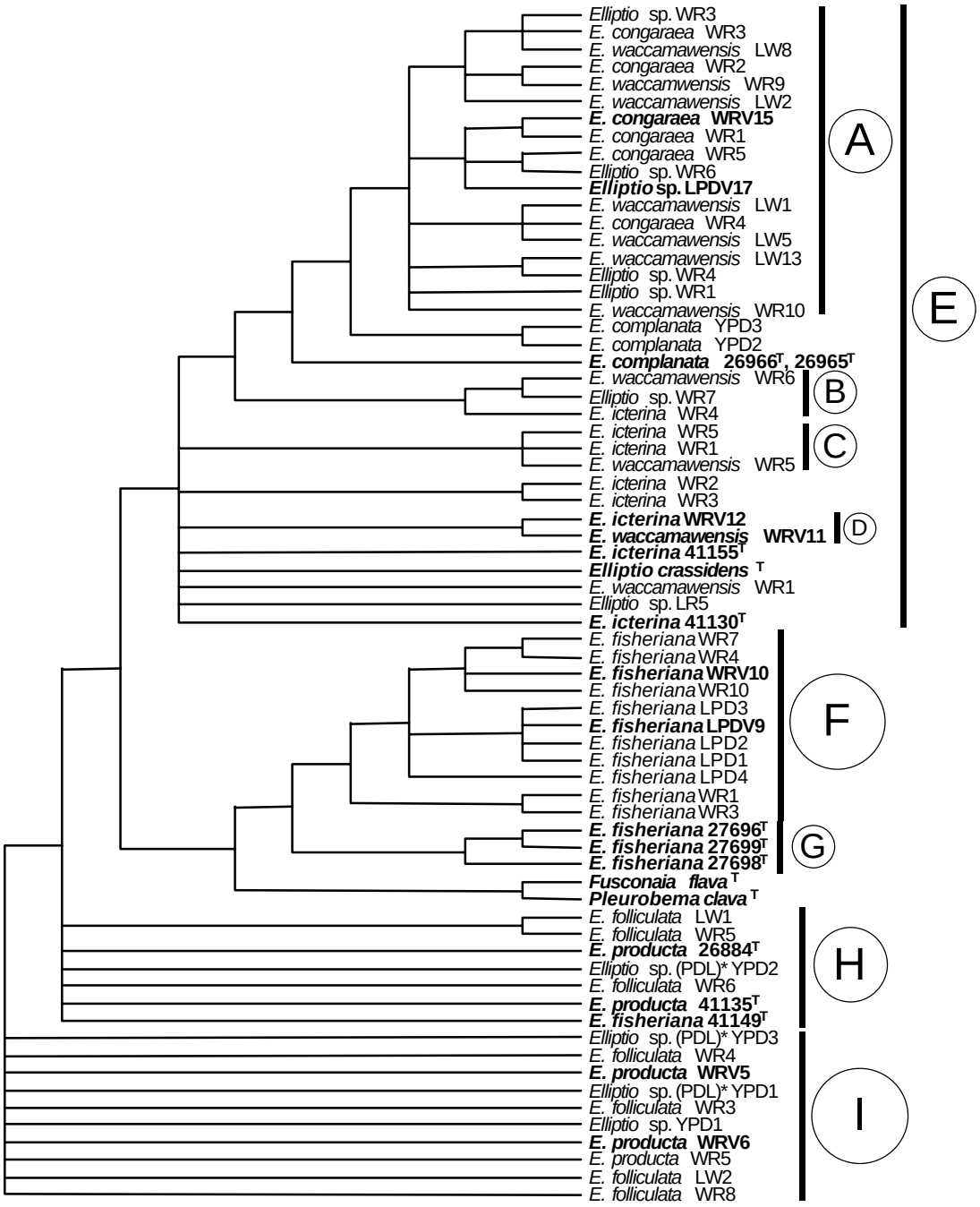

Figure 12. Combined maximum parsimony phylogeny of 16S rRNA gene, *cox1*, and *nad1 Elliptio* sequences. Strict consensus of 447 most parsimonious trees found during partial analysis. Bootstrap values were not able to be calculated and are therefore not shown on the tree. A-I represent important groups within the phylogeny. Taxon label includes species name and sampling location. Topotype sequences and tissue sequences from sacrificed animals are in bold. Topotypes are denoted with a superscript T. Sacrificed animal sequences are denoted with a V preceding the sequence number. * Pee Dee Lance.

The topologies between the *MrBayes* and maximum parsimony trees were fairly consistent; however, there were a few subtle differences. First, similar to the maximum likelihood tree, there was no evidence for the relationship of *Elliptio* sp. (PDL) YPD2 and *E. folliculata* WR6 with the clade containing *E. folliculata* LW1 and *E. folliculata* WR5, which was supported in the *MrBayes* phylogeny by 72% (Figure 10). Also, as with the maximum likelihood analysis, the parsimony analysis grouped *Fusconaia flava* and *Pleurobema clava* reference sequences with the major *E. fisheriana* clade (Figure 12), which was not seen in the MrBayes phylogeny (Figure 10).

Tests of monophyly

The results of The Shimodaira-Hasegawa topology tests (Table 10) demonstrated that the original *MrBayes* consensus tree file for the combined gene dataset was significantly better than the tree created in *MacClade*, which forced *E. waccamawensis* and *L. fullerkati* to each form monophyletic clades. The log likelihood values were significantly (*P* values: 0.018and 0.011, Table 10) greater (less negative) for the *MrBayes* tree than for either the *E. waccamawensis* or the *L. fullerkati* forced monophyly tree. Therefore, the topology tests demonstrated that neither *E. waccamawensis* nor *L. fullerkati* were monophyletic. However, because of the possibility that other relationships within tree were altered during the creation of the forced monophyly clade, a more reliable backbone constraint analysis will be done in *PAUP* to test for monophyly. This method is more reliable because the tree will not be physically altered; instead, the analysis will be constrained so that sequences from the endemic species must group together.

Table 10. Topology test results. This table contains the results from the Shimodaira-Hasegawa (SH) topology tests, which were performed in *PAUP*. The best tree was selected based on higher ln likelihood values (less negative). *P* values less than 0.05 were significant.

Species	**Tree**	**ln Likelihood**	***P* value**
E. waccamawensis	True tree	-11051.54	0.018
	Forced monophyly	-11127.10	
L. fullerkati	True tree	-11051.54	0.011
	Forced monophyly	-11119.27	

DISCUSSION

Comparison of 16S rRNA gene, *cox1*, and *nad1* gene regions

There were advantages and disadvantages of each gene used in this study. The 16S rRNA gene produced the fewest number of characters; however, this gene was the easiest and most successful gene amplified. One of the major disadvantages of 16S rRNA gene was its inability to resolve some of the more recent relationships among *Elliptio* sequences, which will be discussed in more detail below. *cox1* provided more characters than the 16S rRNA gene; however, this gene was the most difficult to amplify. *cox1* was better at solving more recent relationships among *Elliptio* than the 16S rRNA gene, which was also true of *nad1*. *nad1* provided the greatest number of characters and was easier to amplify than *cox1*. However, the *nad1* primers used in this study did not amplify any individuals from the *Uniomerus* clade. With the development of primers more specific for this clade, *nad1* would be as successful as 16S rRNA gene in amplifying freshwater mussel DNA. Overall, there was not a single best gene region for this study. However, *cox1* and *nad1* were more appropriate because they were better than the 16S rRNA gene at resolving recent relationships, which is important in this study because Lake Waccamaw is geologically young (Stager and Cahoon 1987).

Comparison of *MrBayes* phylogenies

All of the datasets produced similar MrBayes phylogenies; however, there were differences within the major clades. For the datasets containing all of the taxa, there were seven major clades. The first clade contained *Elliptio* sequences from *E. congaraea*, *E. waccamawensis*, *E. icterina*, *E. complanata* and the *E. crassidens* reference sequence, which were the *E. complanata*-like forms of *Elliptio* (Davis et al. 1981). In the 16S rRNA gene, there were no internal relationships found in this clade (Figure 6). On the other hand, *cox1*, *nad1*, and

the combined gene phylogenies further divided this major clade into subgroups (Figures 7, 8, and 9, respectively). The second major clade contained all of the lance forms of *Elliptio*. All of the datasets demonstrated that there were two subgroups within this clade. *cox1*, *nad1*, and the combined datasets produced phylogenies in which the lance clade was sister to the *Elliptio* clade discussed above. However, the 16S rRNA gene placed the lance subgroups at the bottom of the *E. complanata*-like *Elliptio* clade. The placement of the lance clade in the 16S rRNA gene phylogeny may have been due to the inability of the 16S rRNA gene to resolve the relationship between these groups of *Elliptio*. The third major clade was the monophyletic clade containing *E. fisheriana*. This clade was consistently resolved by each dataset. The only difference among the datasets was the placement of the *E. fisheriana* topotype sequences. *cox1*, *nad1*, and the combined dataset placed these sequences in a clade, which grouped with the clade containing all of the collected *E. fisheriana* sequences. On the other hand, the *E. fisheriana* topotype sequences in the 16S rRNA gene phylogeny did not form a distinct subclade, but were instead basal to the other *E. fisheriana* sequences, again demonstrating the lower resolution power of the 16S rRNA gene for *Elliptio* relationships.

The next major clades contained sequences from *Lampsilis* species. The larger clade was the main *Lampsilis* clade, which contained *L. fullerkati* and *L. radiata* sequences. All of the datasets produced this same clade, which did not have any internal relationships or subclades. The smaller *Lampsilis* clade contained the cryptic *Lampsilis* species, which was the same for all the phylogenies. The next major clade was monophyletic and contained the three *Leptodea ochracea* sequences from Lake Waccamaw. In the 16S rRNA gene, these sequences were combined in a single haplotype (Haplotype 35 Figure 6). In the 16S rRNA gene, *nad1*, and combined phylogenies, the *L. ochracea* group came out at the base of the *Lampsilis* clade.

However, in the *cox1* phylogeny this clade was found between the main *Lampsilis* clade and the cryptic *Lampsilis* clade. The placement of *Leptodea ochracea* in the *cox1* phylogeny is most likely due to some error within the alignment or the sequences and needs to be re-evaluated. However, this gene is known to produce inconsistent results with freshwater mussel sequences because of the presence of saturation at the third codon position, which was noticed by Graf and Cummings (2006). This saturation resulted in a loss of resolution in the more recent relationships (Graf and Cummings 2006).

The final clade was the *Uniomerus* containing clade found at the bottom of the 16S rRNA gene, *cox1*, and combined gene phylogenies. This clade was not present in the *nad1* phylogeny because none of the *Uniomerus* sequences would amplify at *nad1*. In addition to the *Uniomerus carolinianus* sequences collected from the Lumber and Little Pee Dee River, this clade also contained "*E. complanata*" sequences from the Yadkin/Pee Dee and Lumber Rivers. These "*E. complanata*" individuals were most likely *U. carolinianus* individuals that were misidentified during sample collection. *E. complanata* and *U. carolinianus* are known to be morphologically similar (Karen Lynch, personal communication). Another interesting find is that for all of these analyses, the topotype specimen *E. icterina* NCSM41116 from the Savannah River (Table 5) grouped with the *U. carolinianus* clade as a distinct basal lineage. It is probable that this individual was also misidentified; and its morphology should be rechecked.

Overall, the *MrBayes* phylogenies were similar across each dataset. There were small differences among the datasets, which could be attributed to error in the case of *cox1* and a lack of resolution of more recent relationships in the case of the 16S rRNA gene. Unfortunately, because the dataset was too complicated to run a parsimony analysis, a partition homogeneity test (ILD: Farris et al. 1195) could not be completed in *PAUP* to determine if our three gene

regions were compatible. However, Campbell et al. (2005) were able to run this test with the same three gene regions, which were shown to be compatible. In addition, the usefulness of the partition homogeneity test is questionable and it has been argued the test does not work well for datasets in which portions of the data are evolving at different rates (Yoder et al. 2001), such as the 16S rRNA gene in the present study. The combined gene phylogeny provided a unique insight into the relationships within the separate gene phylogenies by providing higher support for internal clades, in addition to the major clades. It also provided a more complete dataset because relationships of sequences that were not amplified for every gene could still be seen in the combined phylogeny. I was able to get a better picture studying the combined phylogenies than I would have by simply comparing partial datasets. Also, Graf and Cummings (2006) were able to demonstrate that the family Unionidae was monophyletic through the use of a combined analysis, which they were not able to resolve in the single gene analyses. Furthermore, the consistency of the majority of individuals in all of the phylogenies demonstrated that my data was robust and also allowed me to find potential problematic sequences, as in the case of *cox1* *Leptodea ochracea* sequences, which would have been missed if only one gene was sequenced.

Comparison of analyses on the combined *Elliptio* dataset

The small size of the combined *Elliptio* dataset allowed for multiple tree building algorithms to be run on the same dataset to determine if the same topology would be produced for each method. A full maximum likelihood analysis of the data was run; however, because of the nature of the dataset a full maximum parsimony analysis could not be completed. The phylogenies produced by the three tree building algorithms were the same except for a few minor differences in the placement of taxa within major clades. This was especially true for the *Elliptio* lance clade. Overall, there was not a lot of structure within the lance clade for any of the

separate gene analyses except that of *nad1*. The differences in the amount of structure and the lack of support provided by the three different gene regions could be what were causing the discrepancies among the *MrBayes*, maximum likelihood, and maximum parsimony phylogenies of the combined *Elliptio* dataset. In addition to supporting the *MrBayes* topology, the maximum likelihood analysis resulted in bootstrap values which agreed with the *MrBayes* clade probability values. For the most part, the *MrBayes* clade probability values were higher than the maximum likelihood bootstrap support values. This is similar to what was observed by Wilcox et al. (2002) during their comparison of clade probability values and bootstrap support values of their phylogeny. They found that Bayesian values were higher than bootstrap values; however, through simulation experiments they determined that the Bayesian values were more accurate (Wilcox et al. 2002). Alfaro and Holder (2006) noted that if the model and parameters are carefully selected, then the clade probabilities could produce a more enhanced measure of support. Overall, the important result from the comparison of the three analyses produced from different algorithms was that my dataset was robust and continuously produced the same overall phylogeny.

Endemic status of *E. waccamawensis* and *L. fullerkati*

The main take home message from the results of the present study was that neither endemic species was monophyletic. This was seen by the polyphyly of the clades containing *E. waccamawensis* and *L. fullerkati* for each dataset (Figures 6-12), and was further supported by the results of the topology tests (Table 10). Therefore, the status of *E. waccamawensis* and *L. fullerkati* needs to be re-evaluated. Based on the phylogenetic analyses, it looks like *E. waccamawensis* from Lake Waccamaw was derived from an ancestor shared with *E. congaraea*, which originated from the Congaree River in South Carolina, and that the absence of genetic

distinction between *E. congaraea* and *E. waccamawensis* from the lake and river suggest recent separation and/or ongoing gene flow. However, the addition of topotype specimens of *E. congaraea* from the Congaree River is necessary to fully confirm this conclusion. Even though the sacrificed animal sequences for *E. congaraea* grouped with the hemolymph sequences, there needs to be confirmation that these animals represent or are related to the individuals for which the species was described from the Congaree River. Following this, additional studies of gene flow based e. g. on molecular markers could be used to evaluate whether the lake form shows evidence of reproductive isolation.

In addition, my results suggest that *Lampsilis fullerkati* is actually a lake form of *Lampsilis radiata.* There was no genetic differentiation between *L. fullerkati* and *L. radiata found* within the Yadkin/Pee Dee Drainage system. *L. fullerkati* sequences were also related to an individual of *L. radiata* from the Neuse Drainage system. However, to fully evaluate the significance of these results for systematics, the topotype specimens of *L. radiata from* the Potomac River are needed.

The relationships within *Elliptio* from the present study agree with those from Davis et al. (1981), Davis and Fuller (1981), and Davis (1984), and the relationships within *Lampsilis* agree with those from Kat (1983a) and Stiven and Alderman (1992). The genetic studies done by Davis and his colleagues suggested that there are two main groups of *Elliptio* species, those that have lanceolate shell forms and those that relate to the *E. complanata* form, which was confirmed in the present study by the distinct clades formed by each of these groups. The authors also identified the affinity of *E. waccamawensis* individuals with those from the *E. complanata group* (Davis et al. 1981; Davis and Fuller 1981; Davis 1984); this is similar to the relationship seen in clade A of the majority of the phylogenies presented here, in which *E.*

waccamawensis and *E. congaraea* formed a clade that grouped with *E. complanata* sequences. Further, Kat (1983a) and Stiven and Alderman (1992) found evidence that *L. fullerkati* individuals were genetically indistinguishable from *L. radiata* individuals, which was verified by the non-monophyletic clade in each of the present analyses that contained individuals from both presumptive species.

CONCLUSIONS

Overall, the use of multiple gene regions and multiple phylogenetic analyses, demonstrated that the results presented here were consistent and robust. It was determined that the Lake Waccamaw endemic species were not monophyletic, and with the inclusion of topotype specimens from *E. congaraea* and *L. radiata*, questions of the affinities of the Lake Waccamaw endemic freshwater mussel species can be answered. In addition, this study shows that non-lethal hemolymph sampling can be used for studying phylogenetics of freshwater mussels. For every analysis, the tissue sequences from sacrificed animals grouped with the corresponding hemolymph sequences taken from individuals of the same species from the same sampling location. With the use of non-lethal sampling, researchers will be able to obtain more samples, which is necessary to obtain a more accurate understanding of the phylogenetic relationships of unionids (Graf and Cummings 2006). However, there is still the need to include sequences from tissue from a subset of sacrificed animals and topotype specimens to ensure that the animals sampled in the field were not misidentified.

In the future, a more extensive biogeographic study with the use of molecular markers, such as microsatellites, would contribute to the understanding of the origin of the Lake Waccamaw *E. waccamawensis* populations. Samples of *E. congaraea*, *E. complanata*, and *E. waccamawensis* populations collected further downstream the Waccamaw River system, in

addition to the topotype specimens, would provide a more holistic genetic profile of these species throughout their ranges in North and South Carolina. This population assessment could also facilitate conservation biologists in their attempts to restore populations of *E. complanata*.

LITERATURE CITED

Alfaro, M. E. and M. T. Holder. 2006. The posterior and the prior in Bayesian phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*. 37:19-42.

Avise, J. C. 2004. *Molecular markers, natural history, and evolution*. 2nd Ed. Sinauer Associates, Inc. Publishers, Sunderland.

Baker, A. M., C. Bartlett, S. E. Bunn, K. Goudkamp, F. Sheldon, and J. M. Hughes. 2003. Cryptic species and morphological plasticity in long-lived bivalves (Unionoida: Hyriidae) from inland Australia. *Molecular Ecology*. 12:2707-2717.

Berg, D. J., W. R. Haag, S. I. Guttman, and J. B. Sickel. 1995. Mantle biopsy: a technique for nondestructive tissue-sampling of freshwater mussels. *Journal of the North American Benthological Society*. 14(4):577-581.

Belanger, T. V., C. G. Annis Jr., and D. D. VanEpps. 1990. Growth rates of the Asiatic clam, *Corbicula fluminea*, in the Upper and Middle St. Johns River, Florida. *The Nautilus*. 104(1):4-9.

Bieler, R. and P. M. Mikkelsen. 2006. Bivalvia – a look at the branches. *Zoological Journal of the Linnaean Society*. 148:223-235.

Bogan, A. E. 1998. Freshwater molluscan conservation in North America: Problems and practices. *Journal of the Conchology Special Publication No.2*:223-230.

Bogan, A. E. 2002. *Workbook and key to the freshwater bivalves of North Carolina*. North Carolina Freshwater Mussel Conservation Partnership, Raleigh, NC 101 pp, 10 color plates.

Bogan, A. E. 2004. Freshwater bivalves: diversity and distribution of the Unionoida. *J.*

Egypt. Ger. Soc. Zool. 44D:111-120.

Bogan, A. E. and W. R. Hoeh. 2000. On becoming cemented: evolutionary relationships among the genera in the freshwater bivalve family Etheriidae (Bivalvia: Unionoida). Pp.159-168. *In* Harper, E. M., J. D. Taylor, and J. A. Crame, eds. *The evolutionary biology of the Bivalvia*. The Geological Society, London Special Paper No. 177.

Browne, R. A. 1981. Lake as islands: biogeographic distribution, turnover rates, and species composition in the lakes of central New York. *Journal of Biogeography*. 8:75-83.

Brusca, R. C. and G. J. Brusca. 2003. *Invertebrates*. 2nd ed. Sinauer Associates, Inc., Sunderland.

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

Campbell, D. C. 2000. Molecular evidence on the evolution of the Bivalvia. Pp. 31-46 *in* Harper, E. M., J. D. Taylor, and J. A. Crame, eds. *The Evolutionary Biology of the Bivalvia*. The Geological Society of London, London.

Campbell, D. C., J. M. Serb, J. E. Buhay, K. J. Roe, R. L. Minton, and C. Lydeard. 2005. Phylogeny of North American amblemines (Bivalvia, Unionoida): prodigious polyphyly proves pervasive across genera. *Invertebrate Biology*. 124(2):131-164.

Casterlin, M. E., W. W. Reynolds, D. G. Lindquist, and C. G. Yarbrough. 1984.

Algal and physicochemical indicators of eutrophication in a lake harboring endemic species: Lake Waccamaw, North Carolina. *Journal of the Elisha Mitchell Scientific Society* 100(3):83-103.

Cox, C. B., and P. D. Moore. 1993. *Biogeography an ecological and evolutionary approach*. Blackwell Scientific Publications, Oxford.

Curole, J. P., D. W. Foltz, and K. M. Brown. 2004. Extensive allozyme monomorphism in a threatened species of freshwater mussel, *Margaritifera hembeli* Conrad (Bivalvia: Margaritiferidae). *Conservation genetics*. 5:271-278.

Davis, G. M. 1984. Genetic relationships among some North American Unionidae (Bivalvia): Sibling species, convergence, and cladistic relationships. *Malacologia*. 25(2):629-648.

Davis, G. M. and S. L. H. Fuller. 1981. Genetic relationships among recent Unionacea (Bivalvia) of North America. *Malacologia*. 20(2):217-253.

Davis, G. M., W. H. Heard, S. L. H. Fuller, C. Hesterman. 1981. Molluscan genetics and speciation in *Elliptio* and its relationship to other taxa of North American Unionidae (Bivalvia). *Biological Journal of the Linnaean Society*. 15:131-150.

Dowling, T. E., C. Moritz, J. D. Palmer, and L. H. Rieseberg. 1996. Nucleic acids III: Analysis of fragments and restriction sites. *In* Hillis, D. M., C. Moritz, and B. K. Mable. Eds. *Molecular Systematics*, 2nd ed. Sinauer Associates Inc, Sunderland.

Eackles, M. S. and T. L. King. 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.

Evans, B. J., J. Supriatna, N. Andayani, M. I. Setiadi, D. C. Cannatella, and D. J.

Melnick. 2003. Monkeys and toads define areas of endemism on Sulawesi. *Evolution* 57(6):1436-1443.

Farris, J. S., M. Kallersjo, A. G. Kluge, and C. Bult. 1995. Testing significance of incongruence. *Cladistics*. 10:315-319.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*. 3(5):294-299.

Frey, D. G. 1951. The fishes of North Carolina's Bay Lakes and their intraspecific variation. *Journal of the Elisha Mitchell Scientific Society* 67(1):1-44.

Fuller, S. L. H. 1974. Clams and mussels (Mollusca: Bivalvia). Pp. 215-273 *in* C. W. Hart, Jr. and S. L. H. Fuller, eds. *Pollution and ecology of freshwater invertebrates.* Academic Press, New York.

Fuller, S. L. H. 1977. Freshwater and terrestrial mollusks. Pp. 143-194 *in* J. E. Cooper, S. S. Robinson, and J. B. Funderburg, eds. *Endangered and threatened plants and animals of North Carolina*. North Carolina State Museum of Natural History, Raleigh.

Frankham, R. 1998. Inbreeding and extinction: island populations. *Conservation Biology*. 12(3):665-675.

Frankham, R., J. D. Ballou, and D. A. Briscoe. 2002. *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. *Systematic Biology*. 49(4):652-670.

Graf, D. L. 2002. Molecular phylogenetic analysis of two problematic freshwater mussel genera (*Unio* and *Gonidea*) and a re-evaluation of the classification of the Nearctic Unionidae (Bivalvia: Palaeoheterodonta: Unionoida). *Journal of Molluscan Studies*. 68:65-71.

Graf, D. L. and K. S. Cummings. 2006. Palaeoheterodont diversity (Mollusca: Trigonioida + Unionoida): what we know and what we wish we knew about freshwater mussel evolution. *Zoological Journal of the Linnaean Society*. 148:343-394.

Gustafson, L. L., M. K. Stoskopf, A. E. Bogan, W. Showers, T. J. Kwak, S. Hanlon, and J. F. Levine. 2005. Evaluation of a nonlethal technique for hemolymph collection in *Elliptio complanata*, a freshwater bivalve (Mollusca: Unionidae). *Diseases of Aquatic Organisms*. 65:159-165.

Haag, W. R., D. J. Berg, D. W. Garton, and J. L. Farris. 1993. Reduced survival and fitness in native bivalves in response to fouling by the introduced zebra mussel (*Dreissena polymorpha*) in Western Lake Erie. Canadian Journal of Fisheries and *Aquatic Science* 50:13-19.

Hack, J. T. 1969. The area, its geology: Cenozoic development of the southern Appalachians. Pp. 1-17 *in* Holt, P. C., R. L. Hoffman, and C. W. Hart Jr. eds. *The distributional history of the biota of the southern Appalachians*. Virginia Polytechnic Institute, Blacksburg.

Hasegawa, M., H. Kishino, and T. A. Yang. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*. 22:160-174.

Henley, W. F., P. J. Grobler, and R. J. Neves. 2006. Non-invasive method to obtain DNA from freshwater mussels (Bivalvia: Unionidae). *Journal of Shellfish Research*. 25(3):975-977.

Hoeh, W. R., M. B. Black, R. Gustafson, A. E. Bogan, R. A. Lutz, and R. C. Vrijenhoek. 1998a. Testing alternative hypotheses of *Neotrigonia* (Bivalvia: Trigonioida) phylogenetic relationships using cytochrome c oxidase subunit I DNA sequences. *Malacologia*. 40(1-2):267-278.

Hoeh, W. R., A. E. Bogan, K. S. Cummings, and S. I. Guttman. 1998b. Evolutionary relationships among the higher taxa of freshwater mussels (Bivalvia: Unionoida): inferences on phylogeny and character evolution from analyses of DNA sequence data. *Malacological Review*. 31/32(2):123-141.

Hoeh, W. R., A. E. Bogan, and W. H. Heard. 2001. A phylogenetic perspective on the evolution of morphological and reproductive characteristics in the Unionoida. *in* Bauer, G. and K. Wachtler, eds. *Ecology and evolution of the freshwater mussels Unionoida*. Springer-Verlag, Berlin.

Hubbs, C. L. and E. C. Raney. 1946. Endemic fish fauna of Lake Waccamaw, North Carolina. Miscellaneous publications of the Museum of Zoology, University of Michigan, No. 65:1-31.

Huelsenbeck, J.P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.

Hughes, J., A. M. Baker, C. Bartlett, S. Bunn, K. Goudkamp, and J. Somerville. 2004. Past and present patterns of connectivity among populations of four cryptic species of freshwater mussels *Velesunio* spp. (Hyriidae) in central Australia.

Molecular Ecology. 13:3197-3212.

Jenkins, R. E. and W. M. Palmer. 1978. A new species of madtom catfish (Ictaluridae) from the Coastal Plain of the Carolinas. *ASB Bulletin* 25(2):57.

Johnson, R. I. 1970. The systematics and zoogeography of the Unionidae (mollusca: bivalvia) of the southern Atlantic slope region. *Bull. Mus. Comp. Zool*. 140(6): 263-450.

Johnson, R. I. 1984. A new mussel, *Lampsilis* (*Lampsilis*) *fullerkati* (Bivalvia: Unionidae) from Lake Waccamaw, Columbus County, North Carolina, with a list of the other unionid species of the Waccamaw River system. *Occasional Papers on Mollusks* 4(63):305-319.

Kat, P. W. 1983a. Morphologic divergence, genetics, and speciation among *Lampsilis* (Bivalvia: Unionidae). *Journal of Molluscan Studies*. 49:133-145.

Kat, P. W. 1983b. Patterns of electrophoretic and morphologic variability in a widely distributed unionid: an initial survey. *Netherlands Journal of Zoology*. 33(1):21-40.

Kat, P. W. 1983c. Genetic and morphological divergence among nominal species of North American *Anodonta* (Bivalvia: Unionidae). *Malacologia*. 23(2):361-374.

Kat, P. W. 1984. Parasitism and the Unionacea (Bivalvia). *Biological Review* 59:189-207.

Kat, P. W. and G. M. Davis. 1984. Molecular genetics of peripheral populations of Nova Scotian Unionidae (Mollusca: Bivalvia). *Biological Journal of the Linnaean Society*. 22:157-185.

Kelly, M. W. and J. M. Rhymer. 2005. Population genetic structure of a rare unionid

(*Lampsilis cariosa*) in a recently glaciated landscape. *Conservation Genetics*. 6:789-802.

Kerr, J. T. 1997. Species richness, endemism, and the choice of areas for conservation. *Conservation Biology* 11(5):1094-1100.

Kishino, H., T. Miyata, and M. Hasegawa. 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *Journal of Molecular Evolution*. 31:151-160.

Krebs, R. A. 2004. Combining paternally and maternally inherited mitochondrial DNA for analysis of population structure in mussels. *Molecular Ecology*. 13:1701-1705.

Krebs, R. A. R. N. Vlasceanu, and M. J. S. Tevesz. 2003. An analysis of diversity in freshwater mussels (Bivalvia: Unionidae) of the Cuyahoga and Rocky River watersheds (Ohio, USA) based on the 16S rRNA gene. *Journal of Great Lakes Research*. 29(2):307-316.

Lanave, C., G. Preparata, C. Saccone, and G. Serio. 1984. A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution*. 20:86-93.

Lea, I. 1863. Descriptions of twenty-four new species of Unionidae of the United States. *Proceedings of the Academy of Natural Sciences of Philadelphia* 15:191-194.

Leff, L. G., J. L. Burch, and J. V. McArthur. 1990. Spatial distribution, seston removal, and potential competitive interactions of the bivalves *Corbicula fluminea* and *Elliptio complanata*, in a coastal plain stream. *Freshwater Biology* 24:409-416.

Lydeard, C., R. H. Cowie, W. F. Ponder, A. E. Bogan, P. Bouchet, S. A. Clark, K. S. Cummings, T. J. Frest, O. Gargominy, D. G. Herbert, R. Hershler, K. E. Perez, B.

Roth, M. Seddon, E. E. Strong, and F. G. Thompson. 2004. The global decline of Nonmarine mollusks. *Bioscience*. 54(4):321-330.

Lydeard, C., M. Mulvey, and G. M. Davis. 1996. Molecular systematics and evolution of reproductive traits of North American freshwater unionacean mussels (Mollusca: Bivalvia) as inferred from 16S rRNA gene sequences. *Philosophical Transactions of the Royal Society of London B* 351:1593-1603.

MacArthur, R. H. and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. *Evolution*. 17:373-387.

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

Maddison, D. R. and W. P. Maddison. 2000. MacClade 4: Analysis of phylogeny and character evolution. Sinauer Associates, Inc., Sunderland.

Madigan, M. T., J. M. Martinko, and J. Parker. 2003. *Brock biology of microorganisms*. Pearson Education, Inc., Upper Saddle River.

Mayr, E. 1965. Avifauna: turnover on islands. *Science* 150:1587-1588.

McCartney, M. A. and A. E. Wilbur. 2007. An evaluation of hemolymph extraction as a non-lethal sampling method for genetic identification of freshwater mussel species in southeastern North Carolina. Report to the North Carolina Department of Transportation. Pages 1-42.

McMahon, R. F. and A. E. Bogan. 2001. Mollusca: Bivalvia. Pp. 331-429 *in* J. H. Thorp and A. P. Covich, eds. *Ecology and Classification of North American Freshwater Invertebrates*. Academic Press, San Diego.

Myers, N. 2003. Biodiversity hotspots revisited. *Bioscience*. 53(10):916-917.

Myers, N., R. A. Mittermeler, C. G. Mittermeler, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403:853-858.

Nylander, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander, J. A. A. 2006. MrModeltest 2.2 README http://www.csit.fsu.edu/~nylander/mrmodeltest2/mrmodeltest2.html.

Parmalee, P. W. and A. E. Bogan. 1998. *The freshwater mussels of Tennessee*. The University of Tennessee Press, Knoxville.

Peterson, A. T. and A. G. Navarro-Siguenza. 1999. Alternate species concepts as bases for determining priority conservation areas. *Conservation Biology* 13(2):427-431.

Porter, H. J. and K. J. Horn. 1980. Freshwater mussel glochidia from Lake Waccamaw, Columbus County, North Carolina. *The Bulletin of the American Malacological Union, Inc*. pp. 13-17.

Porter, H. J. and K. J. Horn. 1983. Habitat distribution of sympatric populations of selected Lampsiline species (Bivalvia:Unionoida) in the Waccamaw drainage of eastern North and South Carolina. *American Malacological Bulletin*. 1:61-68.

Posada, D. and K. A. Crandall. 1998. Modeltest: Testing the model of DNA substitution. *Bioinformatics*. 14:817-818.

Raley, M. E., J. F. Levine, and A. E. Bogan. 2006. Hemolymph as a nonlethal and minimally invasive source of DNA for molecular systematic studies of freshwater mussels. *Tentacle*. (14):33-34.

Ricciardi, A., R. J. Neves, and J. B. Rasmussen. 1998. Impending extinctions of North

American freshwater mussels (Unionoida) following zebra mussel (*Dreissena polymorpha*) invasion. *Journal of Animal Ecology* 67:613-619.

Roe, K. J. and P. D. Hartfield. 2005. *Hamiota*, a new genus of freshwater mussel (Bivalvia: Unionidae) from the Gulf of Mexico drainages of the southeastern United States. *The Nautilus*. 119(1):1-10.

Roe, K. J. and W. R. Hoeh. 2003. Systematics of freshwater mussels (Bivalvia: Unionoida). Pp. 91-122. *In*: Lydeard, C. and D. R. Lindberg, eds. *Molecular systematics and phylogeography of mollusks.* Smithsonian Books, Washington D.C.

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

Sepkoski, J. J. and M. A. Rex. 1974. Distribution of freshwater mussels: coastal rivers as biogeographic islands. *Systematic Zoology*. 23(2):165-188.

Serb, J. M. and C. Lydeard. 2003. Complete mtDNA sequence of the North American freshwater mussel, *Lampsilis ornata* (Unionidae): An examination of the evolution and phylogenetic utility of mitochondrial genome organization in Bivalvia (Mollusca). *Molecular Biology and Evolution*. 20(11):1854-1866.

Serb, J. M., J. E. Buhay, and C. Lydeard. 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.

Shimodaira, H. and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*.

16(8):1114-1116.

Shute, J. R., P. W. Shute, and D. G. Lindquist. 1981. Fishes of the Waccamaw River Drainage. *Brimleyana* 6:1-24.

Stager, J. C. and L. B. Cahoon. 1987. The age and trophic history of Lake Waccamaw, North Carolina. *Journal of the Elisha Mitchell Scientific Society* 103(1):1-13.

Stiven, A. E. and J. Alderman. 1992. Genetic similarities among certain freshwater mussel populations of the *Lampsilis* genus in North Carolina. *Malacologia*. 34(1-2):355-369.

Swofford, D. L. 2002. Phylogenetic Analysis Using Parsimony (PAUP*). Sinauer, Sunderland, MA.

Tamura, K., and M. Nei. 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.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice.

Tucker, J. K. 1994. Colonization of Unionid Bivalves by the zebra mussel, *Dreissena polymorpha*, in Pool 26 of the Mississippi River. *Journal of Freshwater Ecology* 9(2):129-134.

Turgeon, D. D., J. F. Quinn, Jr., A. E. Bogan, E. V. Coan, F. G. Hochberg, W. G. Lyons, P. M. Mikkelsen, R. J. Neves, C. F. E. Roper, G. Rosenberg, B. Roth, A. Scheltema, F. G. Thompson, M. Vecchione, and J. D. Williams. 1998. *Common*

and scientific names of aquatic invertebrates from the United States and Canada: mollusks. 2nd Edition. American Fisheries Society Special Publication 26: ix-526, Bethesda.

Van der Schalie, H. 1945. The value of mussel distribution in tracing stream confluence. Pap. Michigan Ach. Sci. Arts Letters 20:355-373.

White, L. R., B. A. McPheron, and J. R. Stauffer, Jr. 1994. Identification of freshwater mussel glochidia on host fishes using restriction fragment length polymorphisms. *Molecular Ecology*. 3:183-185.

White, L. R., B. A. McPheron, and J. R. Stauffer, Jr. 1996. Molecular genetic identification tools for the unionids of French Creek, Pennsylvania. *Malacologia*. 38(1-2):181-202.

Wilcox, T. P., D. J. Zwickl, T. A. Heath, and D. M. Hillis. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Molecular Phylogenetics and Evolution*. 25:361-371.

Williams, J. D., M. L. Warren Jr., K. S. Cummings, J. L. Harris, and R. J. Neves. 1993. Conservation status of freshwater mussels of the United States and Canada. *Fisheries*. 18(9):6-22.

Yoder, A. D., J. A. Irwin, and B. A. Payseur. 2001. Failure of the ILD to determine data combinability for slow Loris phylogeny. *Systematic Biology*. 50(3):408-424.

APPENDIX ONE. Table of individuals sequenced for 16S rRNA gene, *cox1*, and *nad1* gene regions. * Pee Dee Lance.

Location	Taxon	16S rRNA gene	*cox1*	*nad1*
Waccamaw River	*Lampsilis* sp. WR1	Y	Y	Y
Waccamaw River	*Lampsilis* sp. WR2	Y	Y	Y
Waccamaw River	*Lampsilis* sp. WR4	Y	Y	Y
Waccamaw River	*Lampsilis* sp. WR5	Y	Y	Y
Waccamaw River	*Lampsilis* sp. WR6	Y	Y	Y
Waccamaw River	***Lampsilis* sp. WRV13**	Y	Y	Y
Waccamaw River	*Elliptio* sp. WR1	Y	Y	Y
Waccamaw River	*Elliptio* sp. WR3	Y	Y	Y
Waccamaw River	*Elliptio* sp. WR4	Y	Y	Y
Waccamaw River	*Elliptio* sp. WR6	Y	Y	Y
Waccamaw River	*Elliptio* sp. WR7	Y	Y	Y
Waccamaw River	*E. waccamawensis* WR1	Y	Y	Y
Waccamaw River	*E. waccamawensis* WR5	Y	Y	Y
Waccamaw River	*E. waccamawensis* WR6	Y	Y	Y
Waccamaw River	*E. waccamawensis* WR9	Y	Y	Y
Waccamaw River	*E. waccamawensis* WR10	Y	Y	Y
Waccamaw River	***E. waccamawensis* WRV11**		Y	Y
Lake Waccamaw	*E. waccamawensis* LW1	Y	Y	Y
Lake Waccamaw	*E. waccamawensis* LW2	Y	Y	Y
Lake Waccamaw	*E. waccamawensis* LW5	Y	Y	Y
Lake Waccamaw	*E. waccamawensis* LW8	Y	Y	Y
Lake Waccamaw	*E. waccamawensis* LW13	Y	Y	Y
Waccamaw River	*E. congaraea* WR1	Y	Y	Y
Waccamaw River	*E. congaraea* WR2	Y	Y	Y
Waccamaw River	*E. congaraea* WR3	Y	Y	Y
Waccamaw River	*E. congaraea* WR4	Y	Y	Y
Waccamaw River	*E. congaraea* WR5	Y	Y	Y
Waccamaw River	***E. congaraea* WRV15**	Y	Y	Y
Yadkin/PeeDee	*E. complanata* YPD1	Y	Y	
Yadkin/PeeDee	*E. complanata* YPD2	Y	Y	Y
Yadkin/PeeDee	*E. complanata* YPD3	Y	Y	Y
Yadkin/PeeDee	*E. complanata* YPD4	Y	Y	
Waccamaw River	*E. icterina* WR1	Y	Y	Y
Waccamaw River	*E. icterina* WR2	Y	Y	Y
Waccamaw River	*E. icterina* WR3	Y	Y	Y
Waccamaw River	*E. icterina* WR4	Y	Y	Y
Waccamaw River	*E. icterina* WR5	Y	Y	Y
Waccamaw River	***E. icterina* WRV12**	Y	Y	Y
Lumber River	*Elliptio* sp. LR1	Y	Y	
Lumber River	*Elliptio* sp. LR2	Y	Y	
Lumber River	*Elliptio* sp. LR3	Y	Y	
Lumber River	*Elliptio* sp. LR4	Y	Y	
Lumber River	*Elliptio* sp. LR5	Y	Y	

Lumber River	***Elliptio* sp. LRV16**	Y	Y	
Waccamaw River	*E. folliculata* WR3	Y	Y	Y
Waccamaw River	*E. folliculata* WR4	Y	Y	
Waccamaw River	*E. folliculata* WR5	Y	Y	Y
Waccamaw River	*E. folliculata* WR6	Y	Y	Y
Waccamaw River	*E. folliculata* WR8	Y	Y	Y
Waccamaw River	*E. producta* WR5	Y	Y	Y
Waccamaw River	***E. producta* WRV5**	Y	Y	
Waccamaw River	***E. producta* WRV6**	Y		Y
Yadkin/PeeDee	*Elliptio* sp. (PDL)* YPD1	Y	Y	Y
Yadkin/PeeDee	*Elliptio* sp. (PDL)* YPD2	Y	Y	Y
Yadkin/PeeDee	*Elliptio* sp. (PDL)* YPD3	Y	Y	Y
Yadkin/PeeDee	*Elliptio* sp. YPD1	Y	Y	Y
Lake Waccamaw	*E. folliculata* LW1	Y	Y	Y
Lake Waccamaw	*E. folliculata* LW2	Y	Y	Y
Waccamaw River	*E. fisheriana* WR1	Y	Y	Y
Waccamaw River	*E. fisheriana* WR3	Y	Y	Y
Waccamaw River	*E. fisheriana* WR4	Y	Y	Y
Waccamaw River	*E. fisheriana* WR7	Y	Y	Y
Waccamaw River	*E. fisheriana* WR10	Y	Y	Y
Waccamaw River	***E. fisheriana* WRV10**	Y		Y
LittlePeeDeeRiver	*E. fisheriana* LPD1	Y	Y	Y
LittlePeeDeeRiver	*E. fisheriana* LPD2	Y	Y	Y
LittlePeeDeeRiver	*E. fisheriana* LPD3	Y	Y	Y
LittlePeeDeeRiver	*E. fisheriana* LPD4	Y	Y	Y
LittlePeeDeeRiver	***E. fisheriana* LPDV9**	Y	Y	Y
Lumber River	*E. complanata* LR1	Y	Y	
Lumber River	*E. complanata* LR3	Y	Y	
Lumber River	*E. complanata* LR4	Y	Y	
Lumber River	***E. complanata* LRV2**	Y	Y	
Lumber River	***E. complanata* LRV3**	Y	Y	
Lake Waccamaw	*L. fullerkati* LW3	Y	Y	Y
Lake Waccamaw	*L. fullerkati* LW10	Y	Y	Y
Lake Waccamaw	*L. fullerkati* LW11	Y	Y	Y
Lake Waccamaw	*L. fullerkati* LW12	Y	Y	Y
Lake Waccamaw	*L. fullerkati* LW15	Y	Y	Y
Big Creek	*L. radiata radiata* BigCreek1	Y	Y	Y
Waccamaw River	*L. radiata radiata* WR1	Y	Y	Y
Waccamaw River	*L. radiata radiata* WR2	Y	Y	Y
Yadkin/PeeDee	*L. radiata* YPD1	Y	Y	Y
Yadkin/PeeDee	*L. radiata* YPD2	Y	Y	Y
Yadkin/PeeDee	*L. radiata* YPD3	Y	Y	Y
Yadkin/PeeDee	*L. radiata* YPD4	Y	Y	Y
Yadkin/PeeDee	*L. radiata* YPD5	Y		Y
Flat River	*L. radiata radiata* FlatRiver1	Y	Y	Y
Lick Creek	*L. radiata radiata* LickCreek1	Y	Y	Y
Tar River	*L. radiata radiata* TarRiver1	Y	Y	Y
ChesapeakeBayBasin/ChesterR.	***E. fisheriana* 27696^{T}**	Y		Y

ChesapeakeBayBasin/ChesterR.	***E. fisheriana* 27698^T**	Y	Y	Y
ChesapeakeBayBasin/ChesterR.	***E. fisheriana* 27699^T**	Y		Y
Savannah Basin/SavannahR.	***E. fisheriana* 41149^T**	Y	Y	Y
Savannah Basin/SavannahR.	***E. icterina* 41155^T**	Y	Y	Y
Savannah Basin/SavannahR.	***E. icterina* 41116^T**	Y	Y	
Savannah Basin/SavannahR.	***E. icterina* 41130^T**	Y	Y	Y
Savannah Basin/SavannahR.	***E. producta* 41135^T**	Y	Y	Y
Savannah Basin/SavannahR.	***E. producta* 26884^T**	Y		
ChesapeakeBayBasin/PotomacR.	***E. complanata* 26965^T**	Y	Y	Y
ChesapeakeBayBasin/PotomacR.	***E. complanata* 26966^T**	Y	Y	Y
Lumber River	*U. carolinianus* LR1	Y	Y	
Lumber River	***U. carolinianus* LRV14**	Y	Y	
LittlePeeDeeRiver	*U. carolinianus* LPD1	Y	Y	
LittlePeeDeeRiver	*U. carolinianus* LPD2	Y	Y	
LittlePeeDeeRiver	***Elliptio* sp. LPDV17**	Y		Y
Lake Waccamaw	*Leptodea ochracea* LW16	Y	Y	Y
Lake Waccamaw	*Leptodea ochracea* LW17	Y	Y	Y
Lake Waccamaw	*Leptodea ochracea* LW18	Y	Y	Y