ABSTRACT

MOLINA, REVERIE ALVAREZ. Morphological and Genetic Description of the Freshwater Mussel, *Elliptio complanata* (Lightfoot, 1786) in the Cape Fear River system, NC. (Under the direction of Jay F. Levine).

The purpose of this research is to provide a preliminary description of the morphological and genetic variation of a cosmopolitan freshwater mussel *E. complanata* from one North Carolina river system, Cape Fear River (CFR). Individuals from CFR were collected and compared with known specimens of *E. complanata* (topotype). Multivariate analyses, such as factor and discriminant analyses were utilized to differentiate the individuals based on thirty morphological shell landmarks. Genetic analyses involved the use of diversity estimates and cluster analyses based on cytochrome oxidase I (COI) sequence and Amplified Fragment Length Polymorphism (AFLP) fingerprint data. Factor analysis suggest that *E. complanata* from CFR maybe differentiated based on the thickness of posterior and anterior shell angles, and obesity of the shells. Significant differences between the CFR samples and topotypes were demonstrated by discriminant analysis of morphological data and by COI gene diversity estimates. This difference corroborated earlier work suggesting geographic delineation of *E. complanata* shell form. Genomic fingerprinting suggests further variation even within the topotypes. Phenotype of the topotypic materials seems to support this genomic variability. Heirarchical cluster analyses of morphometry and genetic data further showed different groups supporting earlier research suggesting high shell form variation within the *E. complanata* species.

MORPHOLOGICAL AND GENETIC DESCRIPTION OF THE FRESHWATER MUSSEL, *ELLIPTIO COMPLANATA* (LIGHTFOOT, 1786) IN THE CAPE FEAR RIVER SYSTEM, N.C.

by

REVERIE ALVAREZ-MOLINA

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

COMPARATIVE BIOMEDICAL SCIENCES

Raleigh

2004

APPROVED BY:

Arthur Bogan	Fred Gould
Wondwossen Gebreyes	Bruce Weir

Jay F. Levine
Chair of Advisory Committee

DEDICATION

For Jon & Isaac Molina

and

my family: Remy, Clemence, Ryea and Raymond Alvarez

BIOGRAPHY

PERSONAL INFORMATION

Name: Reverie Alvarez Molina

Birthdate: 8 June 1972

Husband's name: Jonathan Molina

Child's name: John Isaac Molina

EDUCATION

NORTH CAROLINA STATE UNIVERSITY (NCSU), RALEIGH, NC
Ph.D. Comparative Biomedical Sciences 1999-present
- *Major: Population Medicine, Minor: Biotechnology*
- *Thesis: Morphologic and Genetic Characterization of the freshwater mussel, Elliptio complanata, Advisor: Dr. Jay F. Levine*

UNIV. OF THE PHILIPPINES (UP)-MARINE SCIENCE INSTITUTE, Q.C., PHILIPPINES
39 credites in M.S. Marine Science 1992-1999

MIRIAM COLLEGE, Q. C., PHILIPPINES
B.S. Environmental Planning 1988-1992

RESEARCH/TEACHING/PROFESSIONAL EXPERIENCE

NCSU, RALEIGH, NC
Research Assistant,
Population Health and Pathobiology (PHP), College of Veterinary Medicine *7/01- present*

NCSU, RALEIGH, NC
Teaching Assistant, Biological Sciences Department *8/99-5/01*

UP-MARINE SCIENCE INSTITUTE, QUEZON CITY, PHILIPPINES
Research Associate, Coral Energetics Project *6/94-6/98*
Research Associate, Soft Bottom Ecology Project *8/93-12/96*

ENVIRONMENTAL PRIMEMOVERS OF ASIA, INC, SAN JUAN, METRO MANILA, PHILIPPINES
Research Assistant *5/92-8/93*

AWARDS

- Outstanding Teaching Assistant award, Biological Science Department, NCSU *4/01*
- College Scholar, M.S. Marine Science, College of Science, UP Diliman *1995-1997*
- Area Awardee, B.S. Environmental Planning Department, Miriam College *4/92*
- Dean's Lister, B.S. Environmental Planning Department, Miriam College *1992*

ABSTRACTS

Molina, R.A., J.F. Levine, S.D. Hanlon, T. Savidge and A. Bogan. 2001. A method to monitor the growth of living freshwater mussels. Abstract presented at the 2nd Freshwater Mollusc Conservation Society on 12-14 March at the Westin Convention Center, Pittsburgh, PA.

Nacorda H.M.E., H.T. Yap, ***R.M. Alvarez***, J. Onate-Pacalioga & J. Estacion. 1994. Variations in the structural sediment communities in selected Philippines sites. In C.R. Wilkinson, S. Sudara & L.M. Chou (eds.). Proceedings of the 3rd ASEAN -Australia Symposium on Living Coastal Resources, Chulalongkorn University, Bangkok, Thailand, Vol 1:347-363.

Alvarez R.M., H.T. Yap & H.M.E. Nacorda. 1994. Comparison of two macroinfaunal communities from two embayments in the eastern Philippines. Abstract presented at the 3rd National Symposium in Marine Science on 23-24 May 1994 at the University of the Philippines in the Visayas.

POSTERS

Molina, R.A., M. Raley, A. Bogan and J. F. Levine. 2003. Preliminary results of cytochrome oxidase I (COI) sequencing for the freshwater mussel, *Elliptio complanata* (Lightfoot, 1786). A poster presented at the 2003 College of Veterinary Medicine (CVM) Research Forum held at CVM, North Carolina State University, Raleigh, N.C.

Molina, R.A., M. Raley, A. Bogan and J. F. Levine. 2003. What freshwater mussel tissue can be used for DNA analysis? A poster presented at the 3rd Biennial Symposium of the Freshwater Mollusc Conservation Society on 16 – 19 March 2003 at the Sheraton Imperial Hotel, Durham, NC

PUBLICATIONS

Yap, H.T., ***R. A. Molina***. 2003. Comparison of coral growth and survival under enclosed, semi-natural conditions and in the field. Marine Pollution Bulletin 46: 858-864.

Molina, R., S. Hanlon, T. Savidge, A. Bogan & J. Levine. *In review*. Buoyant weight technique: application to freshwater bivalves. American Malacological Bulletin.

Yap H.T., ***R.M. Alvarez***, H.M. Custodio III and R.M. Dizon. 1998. Physiological and ecological aspects of coral transplantation. Journal of Experimental Marine Biology and Ecology 229: 69-84.

Yap H.T., ***R.M. Alvarez***, H.M. Custodio III, and R.M. Dizon. 1996. Ecological aspects of coral transplantation. Paper presented at the 8th International Coral Reef Symposium, Panama City, Panama, 24-29 June 1996.

ACKNOWLEDGEMENTS

This chapter in my life is a testimony to God's sovereignty over His children. He has led me to the following people who helped me see this dream to fruition:

Dr. Jay F. Levine who has welcomed and given me the chance to develop as a researcher,

Drs. Art Bogan, Wondwossen Gebreyes, Fred Gould and Bruce Weir, my committee members who have guided and shown support and confidence in the work that I was doing,

Dr. Dave O'Malley, Jing Wang, Becky Lough and Tori Batista, for their patience in teaching me the rigors of AFLP fingerprinting,

Dr. Morgan Raley, my laboratory supervisor and friend who unselfishly gave his time and expertise to a sequencing green horn like me,

All my professors, classmates and friends I have made along the way, for making my journey worthwhile,

My family, especially my Mom who halted her life in the Philippines to be physically present with us in Raleigh to help out with Isaac so I could devote time to finish my degree,

My church friends (from the Filipino church, GRACE and CCC) who rallied with me through prayers during my balancing act as a wife, mom and student,

My husband, Jon, who deserves all the praise for his spiritual, emotional, mental and physical support during this crazy but wonderful time,

AND MOST ESPECIALLY TO MY SON, ISAAC, FOR GIVING MORE MEANING TO MY LIFE AND FOR PUTTING MY FEET ON THE GROUND ALWAYS.

TABLE OF CONTENTS

LIST OF TABLES …… ix

LIST OF FIGURES ……... x

1. LITERATURE REVIEW ……... 1

Background ……... 3
Species ……... 3
Taxonomic History and Description ……... 3
Biology and Ecology …… 6
Bivalve Genetic Variation …… 7
Measurement of Genetic Variability …… 9
Study Objectives ……... 11
References …… 11

2. PRELIMINARY EVIDENCE OF MORPHOLOGICAL PLASTICITY IN THE FRESHWATER MUSSEL, *Elliptio complanata* (Lightfoot 1786) BASED ON SHELL MORPHOMETRY …… 17

Abstract …… 17
Introduction …… 18
Methodology …… 19
Study Area …… 19
Sample Collection …… 20
Sample Processing …… 21
Shell Characters …… 21
Morphometric Variables …… 22
Statistical Analyses …… 22
Results …… 23
Discussion …… 25
Acknowledgements …… 27
References …… 28

3. PRELIMINARY EVIDENCE OF MORPHOLOGIC AND GENETIC VARIATION IN THE FRESHWATER MUSSEL, *Elliptio complanata* (Lightfoot 1786) FROM THE CAPE FEAR RIVER, NC …… 48

Abstract …….. 48
Introduction …….. 49
Methodology …….. 50
Sample Location and Collection …….. 50
Morphometric Measurement and Analyses …….. 51
Mitochondrial Sequencing and Data Analysis …….. 52
Correlation Between Morphometry and DNA Sequence …... 54
Results ……... 54
Morphometric Analysis ……... 54
Mitochondrial DNA Analysis ……... 55
General Statistics ……... 55
Phylogenetic Estimates ……... 56
Population Subdivision ……... 56
Correlation of Morphometric Distances and
COI Sequence Similarities ……... 57
Discussion and Conclusion ……... 57
Morphometry ……... 57
COI Sequence ……... 58
Acknowledgement ……... 62
Reference ……... 62

4. GENETIC VARIATION WITHIN THE FRESHWATER MUSSEL, *Elliptio complanata* (Lightfoot, 1786): A PRELIMINARY INVESTIGATION …….. 82

Abstract ……. 82
Introduction ……. 83
Methodology ……. 85
Sample Acquisition ……. 85
AFLP Analysis ……. 86
Analysis of genetic diversity ……. 88
Results ……. 88
Discussion ……. 91
References ……. 94

5. CONCLUSION AND RECOMMENDATIONS ……. 114

6. APPENDICES ……. 121

Appendix A. Protocol for COI sequencing ……. 121
Appendix B. Procedure for AFLP ……. 127

LIST OF TABLES

Table 2.1. Geographic locations of CFR sites …… 30

Table 2.2. *A priori Elliptio complanata* shell forms based on crude descriptions used in the field …… 31

Table 2.3. Twenty-nine shell characters used for morphometric analysis ……. 32

Table 2.4. Summary statistics of 21 morphometric variables used for multivariate analyses ……. 33

Table 2.5. Results of factor analysis performed on morphometric variables that satisfied assumptions of multivariate statistical analyses …….. 38

Table 3.1. Geographic locations of sampled sites ……. 65

Table 3.2. Summary statistics and description of the seven morphometric measurements grouped by site …….. 66

Table 3.3. List of COI haplotypes and corresponding sample numbers ……. 67

Table 3.4. Pairwise comparison of genetic distance for COI sequence data …….. 68

Table 3.5. Population pairwise F_{ST} …….. 71

Table 4.1. Sample population location, description and sample sizes …….. 97

Table 4.2. List of COI haplotypes and corresponding sample numbers …….. 98

Table 4.3. Measures of genetic diversity ……. 99

LIST OF FIGURES

Figure 2.1. Sample site location ……. 39

Figure 2.2. Representative pictures of the different *a priori* shell forms ……. 40

Figure 2.3. Shell outline enclosed by a rectangle and the different points that correspond to a specific shell character measurement based on Table 2.3 …….. 41

Figure 2.4. Orientation of the seven morphometric variables that had high loadings from factor analysis …….. 43

Figure 2.5. Discriminant plot based on site …….. 44

Figure 2.6. Discriminant plot according to shell morphology …….. 45

Figure 2.7. Heirarchical cluster diagram of collected individuals from Cape Fear River drainage …….. 46

Figure 3.1. Sample site location …….. 72

Figure 3.2. Orientation of morphometric variables included in multivariate analyses …….. 73

Figure 3.3. Heirarchical cluster diagram of *Elliptio complanata* collected from Cape Fear River basin, based on seven morphometric variables …….. 74

Figure 3.4. Summary statistics of seven morphometric variables by cluster groups…….. 76

Figure 3.5. Discriminant plot of seven morphometric variables based on sites …….. 78

Figure 3.6. Unrooted, maximum parsimony-based dendrogram of cytochrome oxidase I (COI) sequence data for *Elliptio complanata* ……. 79

Figure 3.7 Genetic diversity indices calculated for the different sites ……. 80

Figure 3.8. Lack of correlation between log-transformed morphometric distance and arc-sine transformed COI sequence similarity ……. 81

Figure 4.1. Site location ……. 100

Figure 4.2. Unrooted, maximum parsimony-based dendrogram of cytochrome oxidase I (COI) sequence data for *Elliptio complanata* ……. 101

Figure 4.3. Dendrograms of the three AFLP primers (1 = first primer: AAG/CAC, 2 = second primer: AAC/CAC, 3 = third primer: ACC/CAT) taken from eight samples of valid *E. complanata* population, which were replicated (a = first replicate, b = second replicate) ……. 102

Figure 4.4. Dendrogram of *E. complanata* based on Pearson correlation and neighbor-joining procedures ……. 104

Figure 4.5. Dendrogram of *E. complanata* based on Pearson correlation and UPGMA procedures ……. 106

Figure 4.6. Dendrogram of *E. complanata* based on Jaccard similarity and neighbor-joining procedures ……. 108

Figure 4.7. Dendrogram of *E. complanata* based on Jaccard similarity and UPGMA procedures ……. 110

Figure 4.8. Average heterozygosity of the different sites, based on three AFLP primers ……. 112

Figure 4.9. Pictures of topotype samples ……. 113

CHAPTER 1

LITERATURE REVIEW

Freshwater mussels (Bivalvia: Unionidae) are filter feeders that siphon phytoplankton and organic detritus suspended or dissolved in the water column (Jorgensen 1966). In this manner they serve as indicators of ecosystem health and help sustain water quality (Goudreau et al. 1993). Like other bivalves, they also serve as food for benthic invertebrates, fish, birds and some mammals (Langdon and Newell 1996). Native Americans consumed and used freshwater mussels for clothing accessories (Parmalee and Bogan 1998). North America possesses the largest diversity of unionid fauna in the world (National Native Mussel Conservation Committee 1998). However, more than 70% of the approximately 300 species are either endangered, threatened or of special concern (Turgeon et al. 1998). This decline, brought about by urban development and agricultural practices, has resulted in direct changes in unionid habitat (Bogan 1993). Because most freshwater mussels are dependent on freshwater fish-hosts for further development, negative changes in the aquatic system not only affect unionids but fish-hosts as well. By modifying natural dispersal and gene flow between populations and reducing effective population size, the factors causing physical and chemical alterations in freshwater mussel habitats may lead to a decline in population genetic diversity (Mulvey et al. 1998). Thus, it is important for conservation strategies to include measures that would protect and maintain not only the physical presence and abundance of

freshwater mussel fish-hosts and populations, but also their genetic characteristics.

Reliable species identification is basic to all ecological, genetic and conservation studies of organisms. Ortmann (1912, 1921) classified unionids based on anatomical, reproductive and developmental characteristics (Hoeh et al. 2001). Other researchers have stressed the importance of shell characters alone or the use of both anatomical and reproductive features to classify freshwater mussels (Davis 1983). These different schemes have led to different and inconsistent classification of unionids (Hoeh et al. 2001), especially for species displaying extreme morphological plasticity such as those from the genus *Elliptio*. Identification of some species within the genus *Elliptio* remains a challenging exercise due to the wide range of shell form variation displayed by individuals in different habitats (Bogan 2002). Failure to resolve freshwater mussel classification inconsistencies could hamper the creation of conservation and/or management measures, which is usually impacted by residential, commercial and municipal development. Through the Endangered Species Act, developers are required to conduct environmental site assessments before obtaining construction permits. At times projects may impact only one or two remaining locations at which a species can be found. Regulatory biologists must often make pragmatic decisions about whether or not the population to be impacted by construction is essential for sustaining a species. To make informed decisions, conservation biologists must have the correct information on the species identity of the impacted population to avoid inadvertent and often costly proposition to development managers. Thus, freshwater mussel biologists continue to look for ways to resolve inconsistencies in species identification and develop a single

uniform scheme for unionid classification.

Advancements in molecular technology have provided new techniques to classify unionids. Davis and Fuller (1981) showed congruence between protein variation and shell and anatomical characters in North American freshwater mussels examined by allozyme electrophoresis. Analysis of mitochondrial DNA (16S ribosomal gene) was also used to reclassify a unionid family (Margaritiferidae) and analyze other North American freshwater mussel taxa (Mulvey et al. 1997). The use of genetic variation paired with different shell and anatomical characters has provided a better classification scheme for unionids (Hoeh et al. 2001). However, the use of molecular genetic tools for classifying freshwater mussels is in its infancy, and the most appropriate techniques for specific conservation needs must be defined.

BACKGROUND

Species

Taxonomic history and description

Elliptio complanata is a freshwater mussel that has a long taxonomic history. The typical *E. complanata* specimen was first reported as *Mya complanata* (Lightfoot, 1786) (Matteson 1948). In 1913, Haas used *Unio violaceus* (Spengler 1793) instead of *M. complanata* (Matteson 1948). Ortmann referred to this species as *Elliptio violaceus* instead of *U. violaceus* (Athearn 1954). Earlier workers referred to this species as *Unio complanatus*

instead of *U. violaceus* based on an unpublished article by Lightfoot's unpublished article. The name *Elliptio complanatus* (Dillwyn, 1817) was assigned to this species by Matteson in 1948 (Matteson 1948). Johnson (1970) recognized over a hundred synonyms for this species (including the ones that are mentioned here) and finally documented this species as *Elliptio complanata* (Lightfoot 1786). Bogan (2002) reported that the type locality for *E. complanata* is in the Potomac River, Washington, D.C. and described the following as a typical shell form:

> "Shell shape trapezoidal to rhomboid or subelliptical, compressed to inflated, shell thickness varies from thin to solid, length 120 mm. Anterior margin is rounded, dorsal and ventral margins are roughly parallel, ventral margin is often straight, posterior margin broadly rounded ending at or near the base in a point or biangulation. Posterior ridge is broad and double and rounded to angular. The posterior slope is flat. Beaks are low and uninflated, beak sculpture consists of 5-6 ridges, the first two or three curved and subconcentric, the rest run parallel to the growth lines, nearly straight in the middle and curved up at both ends. Surface with irregular growth lines and varies from smooth to mat. Left valve has two ragged pseudocardinal teeth and two nearly straight lateral teeth. Right valve has a single pseudocardinal tooth and a single lateral tooth. Interdentum is essentially absent. Beak cavity is shallow. Periostracum is yellowish to brown and blackish, young specimens with indistinct greenish rays present. The rays generally disappear in older shells. Nacre varies from white, pink, salmon to various shades of purple" (Bogan 2002).

Despite the presence of a typical phenotype, *E. complanata* was still recognized to be highly variable in form, which made its identification confusing and difficult. Johnson (1970) recognized that *E. complanata* from large rivers (such as the Neuse, Tar and Roanoke rivers) were subelliptical in shape and had flattened valves while those from small rivers and tributaries were more rhomboid and inflated. In addition, this variability led some authors to

assign subspecies and ecophenotypes. The subspecies, *E. complanatus roanokensis* and *E. complanatus complanatus*, were good examples. The former was reported to be bigger in length, height and diameter than most *E. complanatus* found in the Neuse River, NC and was generally found in large rivers (Walter and Parker 1957, Johnson 1970). The subspecies, *E. complanatus complanatus*, on the other hand, was reported to be a form found in small rivers (Johnson 1970). Synonymous species such as *E. roanokensis* and *E. northamptonensis*, add to the variation present in *E. complanata*. Compared to *E. complanata*, these two species were larger, more elongated, more compressed, subrhomboid in shape and 'usually exhibited a shallow radial depression in front of the posterior ridge' (Athearn 1954). Other freshwater mussels that can easily be confused with *E. complanata* are *E. hopetonensis* (Lea) in Altahama and *E. icterina* (Conrad) and *Uniomerus tetralasmus* in North Carolina (Johnson 1970). Because of the recognized morphological variation present in the Eastern Elliptio, freshwater mussel researchers agreed on this division: ecophenotyes found north of Washington were considered *E. complanata* and those found south of Washington were assigned subspecific names (Johnson 1970). The highly variable shell form of *E. complanata*, particularly in the southeast region of North America, led Bogan (2002) to suggest an *E. complanata* complex consisting of the following species: *E. complanata* (Lightfoot, 1786), *E. congaraea* (Lea, 1831), *E. judithae* Clarke, 1981, *E. roanokensis* (Lea, 1838), *E. steinstansana* Johnson and Clarke, 1983, *E. waccamawensis* (Lea, 1863) and *E. raveneli* (Conrad, 1834).

Biology and Ecology

The genome size for this species had not yet been established. But for the freshwater mussel, *Lampsilis ornata*, its mtDNA gene was 16,060 bp and was comprised of 37 genes (Serb and Lydeard 2003). Karyotype analysis performed by Park and Burch (1995) revealed that *E. complanata*, just like most freshwater mussels that were analyzed in their study, had 38 diploid chromosomes.

The range of Eastern Elliptio had been documented to extend from southeastern Canada to the eastern Gulf drainages, excluding Florida (Mollusc website of Ohio State University - OSU). Its abundance was difficult to establish but limited studies have reported that they were abundant, with a density reaching 15 individuals/m^2 in a Connecticut lake (Dillon 2000). They were the most ubiquitous freshwater mussel species documented during eastern river surveys in North America (Balfour and Smock 1995, Roble and Stevenson 1997, Clayton et al. 2001, Hanson and Locke 2001, Martel et al. 2001). During an ongoing bridge survey along the Cape Fear River (CFR), North Carolina, by NCSU biologists, *E. complanata* were the most abundant freshwater mussel species; more than six forms had already been tentatively classified as *E. complanata* (Chris Eads, personal communication). On the mollusc website of the Ohio State University (OSU), six shell forms had been classified as *E. complanata* and Johnson (1970) recognized the complexity of shell forms apparent in the species *E. complanata* found in North Carolina.

The Eastern Elliptio was found in lentic (slow-flowing) habitats with a broad range in sediment grain size, ranging from sand to fine gravel (Bogan 2002). Its shell form variability

has been related to specific environmental factors, which has been interpreted as morphological plasticity in this species. Hinch and Bailey (1988) suggested that wind exposure was responsible for the shell form differences found in several lakes in south-central Ontario, Canada. The Eastern Ellipito individuals from high wind exposure sites were longer, taller, wider, and had heavier shells. Individuals from low-wind exposure sites displayed the opposite shell form trend. Furthermore, shell height (greatest dorso-ventral dimension) was related to sediment type and water depth. Narrow shells were directly related to coarse sediment and deep water; smaller, obese shells were found in habitats that were shallow and had fine sediments (Hinch et al. 1989). The same authors also reported that shell thickness was also directly related to alkalinity, conductivity and pH of the surrounding water. Individuals found in Canadian lakes that had high alkalinity, conductivity and pH were reported to have thick shells (Hinch et al. 1989).

Bivalve genetic variation

The long-term survival of a species and its ability to respond to environmental changes depends on a population's ability to maintain a certain level of genetic variability (Schonewald-Cox 1983, Booy et al. 2000). For example, high levels of genetic variability accounted for the physiological plasticity of the cockle, *Cerastoderma glaucum Brugiere* (Trotta and Cordisco 1998). In addition, genetic variability in the bivalve, *Macoma balthica*, was found to be negatively correlated with sensitivity to copper (Hummel et al. 1997). In

areas where there were *M. balthica* populations showing increased stress sensitivity, genetic variability was low (Hummel et al. 1997).

Determining genetic variability can also provide a glimpse of a population's genetic structure as well as play an important role in conservation decisions. In studies conducted with freshwater mussels, two patterns of genetic variation were reported. One pattern suggested that some species were largely undifferentiated indicating large, random mating populations and high rates of gene flow occurring in different populations. An example was the *Quadrula quadrula* populations studied by Berg and co-workers (1998) using allozyme frequency. In this study, they found that in large rivers in Ohio, Tennessee and Mississippi, *Q. quadrula* showed little genetic difference among aggregations that are > 1000 km apart (Berg et al. 1998). When freshwater mussels are not genetically undifferentiated, relocation and other conservation efforts can be conducted within a study area without fear of inadvertent genetic exchange.

Despite this finding, a growing number of studies have also documented species isolation and possibly local adaptation. Gene differentiation was apparent in *Elliptio dilatata* surveyed from small streams in Ohio, Tennessee and Mississippi (Berg et al. 1997). Using allozyme frequency differences, Berg and co-workers (1997) found significant allozyme variation between *E. dilatata* aggregations that are < 100 km apart. In the case of *Pyganodon grandis*, Liu and co-workers (1996a) found significant mtDNA variation in specimens collected in different Colorado River drainages. Genetic differences between two disjointed populations of *Potamilus inflatus* were reported by Roe and Lydeard (1997). Johnson and

co-workers (1997, 1998) found significant genetic variation among four different freshwater mussel species in Arkansas. Latitudinal differences in population structure were also reported for *Lasmigona subviridis* from the Atlantic slope using ribosomal and mtDNA sequence variation (King et al. 1999). Based on these studies, undifferentiation of freshwater mussel species could potentially impact conservation measures by restricting or preventing relocation of a species to an area of concern. Villela and co-workers (1998) have suggested that conservation efforts be focused on the protection of existing populations and on discovering and protecting new populations of endangered species, instead of restocking and inadvertently creating new populations that have 'unknown ecological and evolutionary potential'.

Measurement of Genetic Variability

Conservation efforts necessitate that resource managers consider the genetic potential of a specific population. Assessing gene variation within and between mussel populations should be a basic part of every species conservation plan. Molecular analysis of genetic variability and phylogeny of bivalves has been attempted using DNA sequencing of the mitochondrial genes (mtDNA), cytochrome oxidase I (COI,) and/or 16S rDNA (Chase et al. 1998, Dahlgren et al. 2000, Park and Foighil 2000, Hoeh et al. 2001, Stepien et al. 2001). But analyses based on a single gene, such as COI, may provide incomplete information on the species' true population history (Ballard and Whitlock 2004). The inclusion of additional and independent loci form the genome, aside from mtDNA, is necessary to estimate species

tree and to resolve genetic variation within species (Giannasi et al. 2001, Zawko et al. 2001, Ballard and Whitlock 2004). Genetic fingerprinting has become a useful alternative tool, in which, an organism's DNA fingerprints derived from the whole genome, can be documented by specific band patterns or markers (Campbell et al. 1999). Because of the uniqueness of each organism, a specific marker or set of markers can be generated that can differentiate one individual from another. From these sets of markers, biologists can assess the genetic variability present in a population or between populations

The majority of prior freshwater mussel and bivalve studies on genetic variation have been conducted using allozyme electrophoresis (Marsden et al. 1996, Berg et al. 1998, Johnson et al. 1998, Gardner and Thompson 1999). Although time efficient, the quantity of information generated is minimal (Mueller and Wolfenbarger 1999). Microsatellite analysis has improved the quality and quantity of information available for assessing genetic variation within a population, but the cost and time required is prohibitive (Queller et al. 1993, Mueller and Wolfenbarger 1999, Giannasi et al. 2001). Another Polymerase Chain Reaction (PCR)-based technique, amplified fragment length polymorphism (AFLP), has all the qualities of microsatellites, but is markedly less costly (Mueller and Wolfenbarger 1999). Genetic assessment by AFLP has provided functional a relatively inexpensive source of genetic markers in plants (Mariette et al. 2001, Sawkins et al. 2001, Vekemans et al. 2002), vertebrates such as snakes (Giannasi et al. 2001), invertebrates such as mosquitoes (Ravel et al. 2001) and worms (McMichael and Prowell 1999), and bacteria (Janssen et al. 1996).

STUDY OBJECTIVES

In an effort to further define the utility of molecular genetic tools for the classification of freshwater mussels and the genetic variation within populations, combined morphologic shell characters and genetic distinctions were used to describe the common freshwater mussel species, *Elliptio complanata*. The overall objective of the study was to characterize morphological and genetic variation in the freshwater mussel species, the Eastern Elliptio (*Elliptio complanata*, Lightfoot 1786) by examining the different shell forms found in selected sites at the Cape Fear River system in NC and at a site historically described to contain *E. complanata* (topotype site). Specific objectives were:

1) to distinguish and differentiate between different shell forms of *E. complanata* taken from CFR and topotype sites, based on shell morphometry,
2) to identify and assess genetic polymorphisms present among individuals at each site,
3) to determine and compare genetic variability of *E. complanata* individuals, and
4) to compare and contrast the relationship of the sampled population's genetic variation with the sample population's shell form variation.

REFERENCE

Athearn, H.D. 1954. A variation of *Elliptio complanatus* Solander. American Malacological Union News Bulletin and Annual Report, pp. 13-14.

Balfour, D.L. and L.A. Smock. 1995. Distribution, age structure and movements of the freshwater mussel *Elliptio complanata* (Mollusca: Unionidae) in a headwater stream. Journal of Freshwater Ecology 10(3): 255-268.

Ballard, J.W.O. and M.C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729 - 744.

Berg, D.J., W.R. Hoeh, S.I. Guttman. 1997. Alternate models of genetic structure in unionid populations: conservation and management implications. Journal of Shellfish Research 16(1): 322-323.

Berg, D.J., E.G. Cantonwine, W.R. Hoeh and S.I. Guttman. 1998. Genetic structure of *Quadrula quadrula* (Bivalvia: Unionidae): little variation across large distances. Journal of Shellfish Research 17(5): 1365-1373.

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

Bogan, A.E. 2002. Workbook And Key To The Freshwater Bivalves of North Carolina. North Carolina Museum of Natural Sciences, Raleigh, NC, 101 pp, 10 color plates.

Bogan, A.E., J.M. Serb and C. Lydeard. 2003. Determination of the systematic position and relationships of the *Elliptio* complex (Mollusca: Bivalvia: Unionidae) with distributions in North and South Carolina. A Final Report submitted to USGS (Contract 01HQAG0185).

Booy, G., R.J.J. Hendricks, M.J.M. Smulders, J.M. Van Groenendael and B. Vosman. 2000. Genetic diversity and the survival of populations. Plant biol. 2: 379-395.

Campbell, N.A., J.B. Reece, L.G. Mitchell. 1999. Biology, 5th edition. Benjamin Cummings: U.S. 1175 pp.

Chase, M.R., R. J. Etter, M. A. Rex and J. M. Quattro. 1998. Bathymetric patterns of genetic variation in a deep-sea protobranch bivalve, *Deminucula atacellana*. Marine Biology 131(2): 301-308.

Clayton, J.L., C.W. Stihler and J.L. Wallace. 2001. Status of and potential impacts to the freshwater bivalves (Unionidae) in Patterson Creek, West Virginia. Northeastern Naturalist 8(2): 179-188.

Dahlgren, T.G., J.R. Weinberg and K.M. Halanych. 2000. Phylogeography of the ocean quahog (*Arctica islandica*): influences of paleoclimate on genetic diversity and species range. Marine Biology 137(3): 487-495.

Davis, G.M. 1983. Chapter 12: Relative roles of molecular genetics, anatomy and morphometrics and ecology in assessing relationships among North American Unionidae (Bivalvia). *In* G.S. Oxford and D. Rollinson (eds.), Protein Polymorphism: Adaptive and Taxonomic Significance, Systematics Association Special Volume No. 24, Academic Press, London and New York, pp. 193-222.

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

Dillon, Jr., R. T. 2000. The Ecology of Freshwater Molluscs. Cambridge University Press: U.K. 509 pp.

Gardner, J.P.A. and R.J. Thompson. 1999. High levels of shared allozyme polymorphism among strongly differentiated congeneric clams of the genus *Astarte* (Bivalvia: Mollusca). Heredity 82: 89-99.

Giannasi, N., R.S. Thorpe and A. Malhotra. 2001. The use of amplified fragment length polymorphism in determining species trees at fine taxonomic levels: analysis of a medically important snake, *Trimeresurus albolabris*. Molecular Ecology 10: 419-426.

Goudreau, S.E., R.J. Neves, and R.J. Sheehan. 1993. Effects of wastewater treatment plant effluents on freshwater mollusks in the upper Clinch River, Virginia, USA. Hydrobiologia 252: 211-230.

Hanson, J.M. and A. Locke. 2001. Survey of freshwater mussels in the Petitcodiac River drainage, New Brunswick. Canadian Field Naturalist 115(2): 329-340.

Hinch, S.G. and R.C. Bailey. 1988. Within- and among-lake variation in shell morphology of the freshwater clam *Elliptio complanata* (Bivalvia: Unionidae) from south-central Ontario lakes. Hydrobiologia 157: 27-32.

Hinch, S.G., L.J. Kelly and R.H. Green. 1989. Morphological variation of *Elliptio complanata* (Bivalvia: Unionidae) in differing sediments of soft-water lakes exposed to acidic deposition. Can. J. Zool. 67: 1895-1899.

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. Ecological studies 145: 257-280.

Hummel, H., R.H. Bogaards, C. Amiardtriquet, G. Bachelet, M. Desprez, J. Marchand, H. Rybarczyk, B. Sylvand, Y. Dewit, and L. Dewolf. 1997. Uniform variation in genetic traits of a marine bivalve related to starvation, pollution and geographic clines. Journal of Experimental Marine Biology and Ecology 191(2): 133-150.

Janssen, P., R. Coopman, G. Huys, et al. 1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. Microbiology 142: 1881-1883.

Johnson, R.I. 1970. The systematics and zoogeography of the Unionidae (Mollusca: Bivalvia) of the Southern Atlantic Slope region. Bulletin of the Museum of Comparative Zoology 140(6): 264-450.

Johnson, R.L., F.Q. Liang, C.D. Milam and J.L. Farris. 1997. Genetic diversity among several species of unionid mussels in Arkansas. Journal of Shellfish Research 16: 234.

Johnson, R.L., F.Q. Liang, C.D. Milam and J.L. Farris. 1998. Genetic diversity and cellulolytic activity among several species of unionid bivalves in Arkansas. Journal of Shellfish Research 17(5): 1375-1382.

Jorgensen, C.B. 1966. Biology of Suspension Feeders. Pergamon Press, Oxford, 357 pp.

King, T.L., M.E. Eackles, B. Gjetvaj and W. R. Hoeh. 1999. Intraspecific phylogeography of *Lasmigona subviridis* (Bivalvia: Unionidae): conservation implications of range discontinuity. Molecular Ecology 8: S65-S78.

Langdon, C.J. and R.I.E. Newell. 1996. Chapter 6: Digestion and nutrition in larvae and adults. *In* V.S. Kennedy, R.I.E. Newell and A.F. Eble (eds.), The Eastern Oyster, *Crassostrea virginica*, Maryland Sea Grant College, College Park: 231-260.

Liu, H.P., J.B. Mitton and S.K. Wu. 1996a. Paternal mitochondrial DNA differentiation far exceeds maternal mitochondrial DNA and allozyme differentiation in the freshwater mussel, *Anodonta grandis grandis*. Evolution 50: 952-957.

Marriette, S., D. Chagne, C. Lezier, P. Pastuszka, A. Raffin, C. Plomion and A. Kremer. 2001. Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. Heredity 86: 469-479.

Marsden, J.E., A.P. Spidle and B. May. 1996. Review of genetic studies of *Dreissena* spp. American Zoologist 36(3): 259-270.

Martel, A.L., D.A. Pathy, J.B. Madill, C. B. Renaud, S.L. Dean and S.J. Kerr. 2001. Decline and regional extirpation of freshwater mussels (Unionidae) in a small river system invaded by *Dreissena polymorpha*: The Rideau River. Canadian Journal of Zoology 79(12): 2181-2191.

Matteson, M.R. 1948. The taxonomic and distributional history of the fresh-water mussel *Elliptio complanatus* (Dillwyn, 1817). The Nautilus 61(4): 127-132.

McMichael, M. and D.P. Prowell. 1999. Differences in amplified fragment length polymorphisms in fall armyworm (Lepidoptera: Noctuidae) host strains. Annals of the Entomological Society of America 92: 175-181.

Mueller, U.G. and L.L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. TREE 14(10): 389-394.

Mulvey, M., C. Lydeard, D.L. Pyer, K.M. Hicks, J. Brim-Box, J.D. Williams and R.S. Butler. 1997. Conservation genetics of North American freshwater mussels *Amblema* and *Megalonaias*. Conservation Biology 11: 868-878.

Mulvey, M., H. –P. Liu and K.L. Kandl. 1998. Application of molecular genetic markers to conservation of freshwater bivalves. Journal of Shellfish Research 17(5): 1395-1405.

National Native Mussel Conservation Committee. 1998. National strategy for the conservation of native freshwater mussels. Journal of Shellfish Research, 17(5): 1419-1428.

Park, J.K. and D.O. Foighil. 2000. Genetic diversity of Oceanic Island Lasaea (Mollusca: Bivalvia) lineages exceeds that of continental populations in the northwestern Atlantic. Biological Bulletin 198(3): 396-403.

Parmalee, P.W. & A.E. Bogan. 1998. The Freshwater Mussels of Tennessee. The University of Tennessee Press: Knoxville, 328 pp.

Queller, D.C., J.E. Strassmann and C.R. Hughes. 1993. Microsatellites and kinship. TREE 8(8): 285-288.

Ravel, S., N. Monteny, D. Velasco Olmos, J. Escalante Verdugo and G. Cuny. 2001. A preliminary study of the population genetics of *Aedes aegypti* (Diptera: Culicidae) from Mexico using microsatellite and AFLP markers. Acta Tropica 78(3): 241-250.

Roble, S.M. and P.H. Stevenson. 1997. First records of freshwater mussels on the Eastern Shore of Virginia. Banisteria 10:22-24.

Roe, K.J. and C. Lydeard. 1997. Molecular systematics of the freshwater mussel genus *Potamilus* (Bivalvia: Unionidae). Malacologia 39: 195-205.

Sawkins, M.C., B.L. Maass, B.C. Pengelly, H.J. Newbury, B.V. Ford-Lloyd, N. Maxted and R. Smith. 2001. Geographical patterns of genetic variation in two species of *Stylosanthes Sw.* using amplified fragment length polymorphism. Molecular Ecology 10: 1947-1958.

Schonewald-Cox, C.M. 1983. Chapter 25. Conclusions: Guidelines to management: A beginning attempt. In C.M. Schonewald-Cox, S.M. Chambers, B. MacBryde and W. L. Thomas (eds.), Genetic and Conservation, Benjamin/Cummings Co.: USA, pp. 414-445.

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). Mol. Biol. Evol. 20(11): 1854-1866.

Stepien, C.A., B. Morton, K.A. Dabrowska, R.A. Guarnera, T. Radja and B. Radja. 2001. Genetic diversity and evolutionary relationships of the troglodytic 'living fossil' *Congeria kusceri* (Bivalvia: Dreissenidae). Molecular Ecology 10(8): 1873-1879.

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

Trotta, P. and C.A. Cordisco. 1998. Gonadal maturation, conditioning and spawning in the laboratory and maturation cycle in the wild of *Cerastoderma glaucum* Bruguiere. Journal of Shellfish Research 17(4): 919-923.

Turgeon, 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, Second Edition. American Fisheries Society Special Publication, 25: 526 p.

Vekemans, X., T. Beauwens, M. Lemaire and I. Roldan-Ruiz. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. Molecular Ecology 11: 139-151.

Villela, R.F., R.L. King and C.E. Starliper. 1998. Ecological and evolutionary concerns in freshwater bivalve relocation programs. Journal of Shellfish Research 17(5): 1407-1413.

Walter, W.M. and R.A. Parker. 1957. *Elliptio complanatus roanokensis* in the Neuse River. The Nautilus 71(2): 60-63.

Zawko, G., S.L. Krauss, K.W. Dixon and K. Sivasithamparam. 2001. Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). Molecular Ecology 10: 2389-2396.

Website
www.biosci.ohio-state.edu/mooluscs/Elliptio/complanata_group.htm

CHAPTER 2

PRELIMINARY EVIDENCE OF MORPHOLOGICAL PLASTICITY IN THE FRESHWATER MUSSEL, *ELLIPTIO COMPLANATA* (Lightfoot 1786) BASED ON SHELL MORPHOMETRY

R.A. Molina[1], A.E. Bogan[2] and J.F. Levine[1]

[1]Population Health and Pathobiology department, College of Veterinary Medicine, North Carolina State University, 4700 Hillsborough St., Raleigh, NC 27606, Tel. #: (919) 513-6302. Email: rmolina@unity.ncsu.edu, revmolina@yahoo.com, Jay_Levine@ncsu.edu

[2]North Carolina Museum of Natural Sciences, Research Laboratory, 4301 Reedy Creek Road, Raleigh, NC 27607, Tel. #: (919) 733-7450 ext 753. Email: arthur.bogan@ncmail.net

ABSTRACT

Purported specimens of *Elliptio complanata* from the Cape Fear River (CFR) system in North Carolina were collected for shell form examination and description using morphometry. Thirty morphometric variables were assigned on each individual shell using the measurement option of ArcView GIS Version 3.2a. Factor, discriminant and cluster analyses were utilized to analyze the individuals based on the different shell landmarks. Factor analysis revealed that the collected individuals maybe distinguished based on the width of the posterior and anterior angles, length of the dorsal anterior side and obesity of the shells. Discriminant analysis showed that *E. complanata* from the CFR is significantly different from topotypic specimens giving credence to the geographic delineation of this species, as suggested by earlier researchers. Heirarchical cluster analysis suggested that there may possibly be three *E. complanata*-like groups, aside from the 'true' *E. complanata* (or topotypes).

INTRODUCTION

North America possesses the highest diversity of freshwater mussels with close to 300 recognized unionids (Turgeon et al. 1998). The majority of unionids are found in the eastern United States, particularly the Ohio, Tennessee and Coosa-Alabama river drainages (Davis and Fuller 1981). In the Atlantic slope region, *Elliptio complanata* is recognized as the most widely distributed and abundant freshwater mussel (Johnson 1970). Its wide distribution and abundance has been attributed to its adaptability to thrive in varying environmental conditions. Bogan (2002) reported that *E. complanata* is found in both slow and fast-moving bodies of water, regardless of sediment type. The apparent lack of habitat preference of *E. complanata* has contributed to its ability to manifest variable shell forms (Athearn 1954). The association of these shell forms with habitat differences suggests that *E. complanata* displays marked morphological plasticity (Matteson 1948; Hinch and Bailey 1988; Hinch et al. 1989).

Biologists have recognized for decades the morphologic variation displayed by *E. complanata* (Matteson 1948; Athearn 1954; Johnson 1970) and its identification remains confusing and difficult. Earlier workers agreed that ecophenotypes found north of Washington, D.C. are typical *E. complanata* species while those found south of it are subspecies (Matteson 1948). Johnson (1970) grouped *E. complanata* of various forms under one species. Kat reported in the 1980s (1984) *E. complanata* discrimination from a location farther south of Washington. Based on electrophoretic and morphological analyses, Kat (1984) argued that *E. complanata* from NE Maryland are similar to those in Nova Scotia,

while those from SE Maryland and Delaware are more similar to populations from Virginia and North Carolina. Bogan and co-workers (2003) reported that morphologic variation in the genus *Elliptio* is most extensive in North and South Carolina. Bogan (2002) suggested that *E. complanata* in North Carolina represent a complex of *Elliptio* that seem to intergrade between and among: *E. complanata* (Lightfoot, 1786), *E. congaraea* (Lea, 1831), *E. judithae* Clarke, 1981, *E. roanokensis* (Lea, 1838), *E. steinstansana* Johnson and Clarke, 1983, *E. waccamawensis* (Lea, 1863) and *E. raveneli* (Conrad, 1834). Exact quantification of these observations, however, is needed to clarify the *Elliptio* species complex in NC. Accordingly, this study was initiated to describe and examine the morphological variation present in *E. complanata*. Morphologic distinctions were made by comparing samples found at a river system in North Carolina with topotypic specimens (those collected from their type locality), using shell morphometry.

METHODOLOGY

Study area

In the Cape Fear River (CFR) basin, tributaries of the middle section of the Deep River were chosen from Polecat Creek in Guilford and Randolph County downstream to the Bear Creek watershed in Moore County (Figure 2.1). This region is 1570.29 km^2 in total area and covers portions of Guilford, Randolph, Chatham, Moore and Montgomery Counties in North Carolina. Asheboro and Ramseur are the main municipalities, but the outskirts of Greensboro (Pleasant Garden area) lie within the uppermost portions of the Polecat Creek watershed. This part of the CFR basin has a variety of geologic formations and streambed types ranging from bedrock and boulder to very sandy.

We selected mussel populations at thirty-eight sites in the CFR basin and selected six sites based on the sites with the most abundant *Elliptio complanata* populations (Figure 2.1). A seventh site, Mary's Creek, was also selected. Mary's Creek is a tributary to the Haw River in southeastern Alamance County within the CFR basin.

Sample collection

A total of twenty individuals, displaying different shell forms, were collected from each of the six CFR sites (Table 2.1). The number of samples collected at specific sites was dictated by the collection permits approved by the North Carolina Wildlife Resource Commission. Samples were sorted into nine *a priori* shell forms that were based on crude descriptions as encountered in the field (Table 2.2, Figure 2.2). Sample collection of a specific morphotype depended on the number of dominant shell morphologies present in an area. For example, if there were four dominant shell forms for *E. complanata* in one area, five individuals from each shell morph were collected; if there were only two, then ten individuals from each shell type were collected.

Only seven samples were taken from one of the CFR sites (MC) due to permitting constraints. Five topotypic specimens (those collected from the type locality: Potomac River, near Washington, D.C.) were provided by the North Carolina Museum of Natural Sciences. Topotypic specimens were classified as 'typical' *E. complanata* based on the following literature description of external shell characters and was thus assigned as shell form one:

"Shell shape trapezoidal to rhomboid or subelliptical, compressed to inflated, shell thickness varies from thin to solid, length 120 mm. Anterior margin is rounded, dorsal and ventral margins are roughly parallel, ventral margin is often straight, posterior margin broadly rounded ending at or near the base in appoint or biangulation. Posterior ridge is

broad and double and rounded to angular. The posterior slope is flat. Beaks are low and uninflated, beak sculpture consists of 5-6 ridges, the first two or three curved and subconcentric, the rest run parallel to the growth lines, nearly straight in the middle and curved up at both ends (Bogan 2002)".

Shell form one served as 'control' upon which all other shell forms were compared and/or contrasted. The other shell forms are described in Table 2.2.

Sample processing

Mussels were collected by hand, between May and June 2002. Sampled mussels were placed in net bags, which were later wrapped in wet burlap cloth, kept in coolers and transported to the laboratory alive. Crude shell forms were etched on individual shells for easier identification in the laboratory where each individual was cleaned and tagged.

Shell characters

In the laboratory, digital pictures of individual animals with two cm ruler were taken. The digital camera, which was arranged such that it was facing the shell subjects, was set 1 foot away from individual samples which were laid flat on a level surface. Morphometric measurements were then taken from the digital pictures using the measure tool of ArcView GIS Version 3.2a.. A rectangle was drawn around the shell and its horizontal and vertical departure from specific points on the shell, were measured (Figure 2.3). There were a total of 29 shell characters; three of which were the traditional morphometric measures of shell length, height and width; 26 were modified measurements from Davis (1983) (Table 2.3).

Distance measurements from ArcView were tabulated and then converted to metric measurements. This was accomplished by dividing tabulated values with the distance

measure of one centimeter. Measurement of some individuals was done twice to establish intra-observer efficiency and for these individuals, average of two measurements was taken.

Morphometric variables

Based on the 29 shell characters, we generated 30 morphometric variables for multivariate analyses. To rule out non-uniformity of values, raw measures of length (L), height (H) and width (W) were not included in the analyses but were instead used to derive ratios for twenty-two shell characters: characters 3 to 15 were divided by height measurement, and characters 16 to 24 were divided by shell length. Characters 25 to 28, which were angle measurements were used as is for the analyses. Proportions of L/W, L/H and H/W were also included as variables to represent shell obesity. An additional variable, shell volume (LxHxW) was added to complete 30 morphometric variables for multivariate statistical analyses.

Statistical analyses

Prior to multivariate analyses, assumption of multivariate normality was calculated following the method suggested by Johnson and Wichern (2002). Data transformation was performed to satisfy assumptions of multivariate normality. Tukey's pairwise combination of the 30 morphometric variables was also tested to determine linear combinations of the different variables based on site and *a priori* shell form. Then, metric measurements of the 30 morphometric variables were used to perform factor, discriminant and cluster analyses. Factor analysis is a multivariate statistical method that generates estimates for the population parameters of the observed dataset (Johnson and Wichern 2002). As such, it was used to determine the sufficient number of factors that could explain the data and determine which of the 30 morphometric variables are useful in differentiating the samples. This analysis was

performed using 'Factor Analysis' under the 'Multivariate' option found in the program, MINITAB for Windows (Release 12).

Discriminant analysis is a multivariate statistical tool that separates a dataset based on discriminants (Johnson and Wichern 2002). This analysis was utilized to separate the collected *E. complanata* individuals according to site and/or *a priori* shell form. Only the variables that showed significant linear combinations between site and shell morphology, were utilized for the analysis. The 'Discriminant' option of the 'Multivariate Method', under the 'Analyze' option in SAS-JMP (release 5.0) was utilized to perform this analysis.

Cluster analysis was utilized to determine the probable number of *E. complanata* 'species groups' based on the present dataset. This analysis was performed using the 'Cluster' command, under the 'Multivariate' option of the program MINITAB.

RESULTS

Different shell characters were found in each collection site (Table 2.3). Sites one and four display six shell forms (shell forms one to four, six, eight and forms one to three, five, seven, eight, respectively). Three sites displayed five shell forms: Site two – Forms one, two, four, six, eight; Site five – Forms one to four, seven; Site six – Forms one to five. Site three had the shell forms one, two, three and nine; while site seven only had two forms (one and six). Site eight, source of the topotypic specimens, had only one shell form (form one), which is the typical or classic *E. complanata* shell morphology.

Of the 30 morphometric variables, only 21 variables satisfied assumptions of multivariate normality and pairwise linearity (Table 2.4). Variables 4, 7, 11, 12, 15, 16, 17, 24 and 28 were removed because of their failure to satisfy statistical assumptions despite

transformation. So, succeeding multivariate analyses were conducted only on the 21 morphometric measurements.

Factor analysis explained 41.6% of the variability when two unrotated and rotated factors were used but percent variability of unrotated factors were more balanced (unrotated: factor one = 21.2%, factor two = 20.5%; rotated: factor one = 37%, factor two = 4.7%). So, preference was given to the unrotated factors in explaining the different variables.

With the unrotated factors, factor one had the lowest loadings on two of the variables that had highest loadings in factor two (L/H and logL/W = measures of obesity) (Table 5). The other two variables that had high loadings in factor two were angle measures: 26 and 27 or posterior and anterior angles, respectively. The first three highest loadings in factor 1 were 14/H, 25 and 6/H, respectively. The orientation of these variables on *E. complanata* shell described measures of obesity and posterior and anterior angle measurements (Figure 2.4).

Nineteen percent of the samples were misclassified by discriminant analysis when site was assigned as the discriminant. Overlapping ellipses (pertaining to 95% confidence limit) in the canonical plot refer to significant similarities among sites (Figure 2.5). The plot was able to separate out site 8 (Washington, D.C.) from the North Carolina sites (sites one to six: CFR; site seven: MC).

Discrimination by *a priori* shell form misclassified 29% of the data. The canonical plot revealed similarities among forms one, two, three, seven, and eight based on the overlapping ellipses (Figure 2.6). Significant differences between shell forms four, six and five were revealed by non-intersecting ellipses. Shell form nine was shown to be similar to all the *a priori* shell forms because its ellipse intersected with the ellipses of all the other shell forms.

Cluster analysis divided the dataset into two major groups (Figure 2.7). Group I consisted of four sub-groups; Group II had three sub-groups. Individuals from site two all fell within the first major group. Each sub-group was not discernible by either site or form pattern, except for one sub-group, D under group I, which was predominantly comprised of all the topotypic specimens.

DISCUSSION

There appear to be four distinct *E. complanata* shell forms in the Cape Fear River (NC) region, based on cluster analysis of morphometry data. Fuller (1972) hypothesized that the Cape Fear River (CFR) is an area that encouraged unionid speciation due to isolation, which resulted from the inundation of interglacial seas. The timing of this event may have coincided with the probable Pleistocene marine invasions of the Carolinian Coastal Plain. During this time, the upper CFR system was blocked from the Coastal Plain fresh water by the interglacial seas, except for small portions of the Neuse and Peedee systems. Despite this, he further pointed out that this hypothesis had been in dispute due to lack of marine fossil evidence from deposits in this area.

Discrimination plot of *a priori* shell form shows overlapping confidence intervals, which suggest similarities of the different shell forms. This suggests that *a priori* shell forms can better be distinguished using descriptions other than the nine crude characters used at the beginning of the study. Better descriptors for the different shell forms maybe found in results from factor analysis, which suggested that measures of obesity and shell angles seem to differentiate the collected *E. complanata* individuals. Factor analysis yielded high loadings on variables 14, 6 and 25 (factor one) as well as variables L/H, logL/W, 26 and 27 (factor

two). Variables 6 (dorsal anterior tangent) and 14 (dorsal departure from anterior vertical congruence) were measurements relating to the anterior angles. Angle 25 (vertical distance from anterior vertical contact to #21 - ventral anterior departure) is also related to the anterior angle. Angles 26 and 27 are the posterior and anterior angles, respectively. Variables L/H and logL/W are measures of obesity (as noted above). Inspection of the individual shells seem to suggest that the collected individuals of *E. complanata* could be differentiated based on the thickness of posterior and anterior angle, length of the dorsal anterior side of the shell, and shell obesity.

Site discrimination suggests that North Carolina *E. complanata* are different from what is typically identified as *E. complanata* (Figure 5). This finding seems to be consistent with observations made by earlier researchers regarding the apparent geographic delineation of *E. complanata* – individuals found north of Washington have shell characteristics that are close to the original specimen described as *E. complanata*; those that are found south of Washington maybe considered subspecies due to varying degrees of shell morphology that they display (Matteson 1948).

Environmental differences maybe one reason for the occurrence of varied shell forms for this species. Environmental differences in specific habitats where *E. complanata* individuals were found were suggested to be agents for causing varied shell forms. These characteristics include wind exposure, sediment type, water depth and water chemistry, such as alkalinity, conductivity and pH and are suggested to affect the length, obesity and thickness of *E. complanata* shells (Hinch and Bailey 1988; Hinch et al. 1989).

We identified *E. complanata* based on an individual's morphometric similarity with topotypic specimens. Using this definition, results from cluster analysis suggest that only

three individuals can be considered *E. complanata*. These individuals grouped with all the topotypic specimens under sub-group D (group I). The remaining three sub-groups within group I maybe considered *E. complanata*-like because of their grouping with the topotypes. Group II maybe considered misidentified *E. complanata* individuals because of their separation away from the earlier group (Figure 2.7).

To conclude, we suggest that the collected *E. complanata* from CFR maybe differentiated using landmark measures of shell obesity and angle measures reported in this study. Site differences from factor analysis show how distinct the topotypes are from the *E. complanata* specimens we collected. Heirarchical clustering suggests the presence of four species-group displaying morphometric measurements that are closely similar to *E. complanata*. Although these are novel findings, they still need to be corroborated with additional measures of species identification such as internal anatomical examination and molecular identification. These additional identification measures need to include other *E. complanata* specimens from other locations and *Elliptio* species for comparison with the present results.

ACKNOWLEDGEMENTS

The authors would like to thank the following for their invaluable assistance: Chris Eads and all the biologists of the Freshwater Mussel Research team at the College of Veterinary Medicine (NCSU), Morgan Raley (NCSU-CVM and NCSM), Dr. Peter Bloomfield (NCSU-Statistics), Judith Johnson-Ratcliff (NCWRC) and Jonathan Molina (MPO-RSMAS/UM). This work is dedicated to J. Isaac Molina.

REFERENCES

Athearn, H.D. 1954. A variation of *Elliptio complanatus* Solander. American Malacological Union News Bulletin and Annual Report: 13-14.

Bogan, A.E. 2002. Workbook and key to the freshwater bivalves of North Carolina. North Carolina Museum of Natural Sciences, Raleigh, NC. 101 pp. 10 color plates.

Bogan, A.E., J. M. Serb, and C. Lydeard. 2003. Determination of the systematic position and relationships of the *Elliptio* complex (Mollusca: Bivalvia: Unionidae) with distributions in North and South Carolina. Final report submitted to USGS (Contract # 01HQAG0185). 45 pp.

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. 1983. Chapter 12. Relative Roles of Molecular Genetics, Anatomy, Morphometrics and Ecology in Assessing Relationships Among North American Unionidae (Bivalvia). Pp. 193-222, *In* Oxford, G.S., and D. Rollinson (Eds). Protein Polymorphism: Adaptive and Taxonomic Significance. Systematics Association Special Vol (24), Academic Press, London & NY.

Fuller, S.L.H. 1972. *Elliptio marsupiobesa*, a new freshwater mussel (Mollusca: Bivalvia: Unionidae) from the Cape Fear River, North Carolina. Proceedings of the Academy of Natural Sciences of Philadelphia 124(1): 1-10.

Hinch, S.G., and R.C. Bailey. 1988. Within- and among-lake variation in shell morphology of the freshwater clam *Elliptio complanata* (Bivalvia: Unionidae) from south-central Ontario lakes. Hydrobiologia 157: 27-32.

Hinch, S.G., L. J. Kelly, and R.H. Green. 1989. Morphological variation of *Elliptio complanata* (Bivalvia: Unionidae) in differing sediments of soft-water lakes exposed to acidic deposition. Canadian Journal of Zoology 67: 1895 – 1899.

Johnson, R.I. 1970. The systematics and zoogeography of the Unionidae (Mollusca: Bivalvia) of the Southern Atlantic Slope Region. Bulletin of the Museum of Comparative Zoology 140(6): 264-450.

Johnson, R.A. and D.W. Wichern. 2002. Applied Multivariate Statistical Analysis, Fifth Edition. Prentice Hall, New York, NY. 767 p.

Kat, P. W. 1984. Molecular genetics of peripheral populations of Nova Scotian Unionidae (Mollusca: Bivalvia). Biological Journal of the Linnean Society 22: 157-185.

Matteson, M. R. 1948. The taxonomic and distributional history of the fresh-water mussel *Elliptio complanatus* (Dillwyn, 1817). The Nautilus 61(4): 127-132.

Matteson, M.R. 1948. Life history of *Elliptio complanatus* (Dillwyn, 1817). American Midland Naturalist 40(3): 690-723.

Turgeon, 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, Second Edition. American Fisheries Society Special Publication, 25: 526 p.

Table 2.1. Geographic locations of CFR sites. Numbers of individuals used in analyses are in parenthesis. Legend for dominant substrate: Clay = 0, Silt = 1, Sand = 2, Gravel = 3, Cobble = 4, Boulder = 5, Bedrock = 6.

Counties	Latitude	Longitude	Stream	Site number	Site name	Dominant substrate	Specimen number
Randolph	35.65	-79.78	Richland Creek (upstream)	1 (n = 20)	Richland1	4.3	379 to 398
Randolph	35.79	-79.67	Sandy Creek	2 (n = 19)	Sandy	4.4	399 to 418
Guilford	35.92	-79.80	Polecat Creek	3 (n = 20)	Polecat	3.6	419 to 438
Randolph	35.61	-79.68	Richland Creek (downstream)	4 (n = 20)	Richland2	4.5	439 to 458
Randolph	35.53	-79.64	Fork Creek	5 (n = 20)	Fork	4.8	459 to 478
Randolph	35.60	-79.58	Brush Creek	6 (n = 19)	Brush	5.4	479 to 498
Alamance			Mary's Creek	7 (n = 7)	Mary	2	499 to 505
Washington, D.C.	39.00	77.25	Potomac River	8 (n = 5)	Topotype		506 to 510

Table 2.2. *A priori Elliptio complanata* shell forms based on crude descriptions used in the field.

Shell forms	Description
1	Typical *E. complanata* form as described in Bogan (2002).
2	Fatter & shorter compared to form 1; L/W is smaller and umbo is more prominent than form 1.
3	L/W is smaller thus giving it a fatter shape and umbo is more prominent than form 2; Hinge is also flatter than form 2's.
4	L/W is bigger than form 3.
5	L/W is bigger giving it a fatter appearance than form 4.
6	Elongated form.
7	Looks like form 3 but is bigger and fatter; Shells are highly eroded.
8	Looks like form 1 but L/W is bigger and more prominent.
9	Looks like *Strophitus sp.*

Table 2.3. Twenty-nine shell characters used for morphometric analysis. Characters with asterisk (*) were not included in multivariate statistical analyses due to their failure to satisfy assumptions of multivariate normality.

Traditional shell measurements	
Shell length	1
Shell height	2
Shell width	29
Height measurements	
1/2 distance from umbo to anterior end; ventral side	3
* 1/2 distance from umbo to anterior end; dorsal side	4
Ventral anterior tangent	5
Dorsal anterior tangent	6
* 2/3 distance from umbo to posterior end; ventral	7
2/3 distance from umbo to posterior end; dorsal	8
Ventral posterior tangent	9
Dorsal posterior tangent	10
* Posterior vertical contact	11
* Anterior vertical contact	12
Dorsal departure from posterior vertical congruence	13
Dorsal departure from anterior vertical congruence	14
* Height from center of ventral concavity	15
Length measurements	
* Dorsal line from umbo	16
* Hinge line	17
Dorsal congruence	18
Ventral congruence	19
Ventral posterior departure	20
Ventral anterior departure	21
Distance from umbo to anterior end	22
Dorsal posterior departure	23
* Length from center of ventral concavity to posterior end; Scored "0" if no concavity	24
Angle measurements	
Vertical distance from anterior vertical contact to #21	25
Posterior angle	26
Anterior angle	27
* Angle departing from hinge line	28

Table 2.4. Summary statistics of 21 morphometric variables used for multivariate analyses. Numbers of morphometric variables correspond to assigned values indicated in Table 2.3. Vol = shell volume or LxHxW; L = shell length; H = shell height; W = shell width.

			Morphometric variables																			
	Shell		**Values divided by H**							**Values divided by L**						**Angle measurement**							
Site	**forms**		**3**	**5**	**6**	**8**	**9**	**10**	**14**	**18**	**19**	**20**	**21**	**22**	**23**	**25**	**26**	**27**	**28**	**Vol**	**L/H**	**L/W**	**H/W**
1	**1**	**Avg**	0.11	0.28	0.35	0.14	0.17	0.43	0.55	0.47	0.13	0.44	0.23	0.32	0.27	0.44	1.99	31.96	46.06	68.07	1.70	3.01	1.77
(n = 20)	(n = 3)	**Std**	0.01	0.03	0.07	0.04	0.05	0.10	0.12	0.07	0.01	0.14	0.10	0.11	0.05	0.05	0.31	1.81	6.22	14.02	0.01	0.18	0.11
	2	**Avg**	0.11	0.26	0.35	0.14	0.21	0.51	0.62	0.50	0.13	0.53	0.16	0.31	0.30	0.43	2.08	34.31	44.86	80.80	1.62	2.77	1.71
	(n = 5)	**Std**	0.04	0.04	0.04	0.05	0.02	0.03	0.05	0.04	0.02	0.03	0.01	0.04	0.04	0.04	0.17	3.43	3.17	10.78	0.05	0.07	0.05
	3	**Avg**	0.12	0.27	0.30	0.10	0.09	0.44	0.56	0.44	0.15	0.51	0.15	0.34	0.28	0.41	1.81	31.97	42.26	61.67	1.77	3.28	1.85
	(n = 2)	**Std**	0.01	0.04	0.04	0.03	0.01	0.05	0.04	0.04	0.05	0.02	0.01	0.01	0.02	0.02	0.16	0.92	0.30	8.03	0.06	0.18	0.04
	4	**Avg**	0.11	0.26	0.30	0.12	0.14	0.42	0.52	0.39	0.17	0.45	0.15	0.31	0.28	0.40	1.64	30.34	39.50	47.50	1.78	3.15	1.77
	(n = 2)	**Std**	0.02	0.02	0.01	0.03	0.04	0.05	0.08	0.01	0.01	0.00	0.05	0.08	0.00	0.07	0.03	4.83	0.08	5.16	0.02	0.01	0.02
	6	**Avg**	0.13	0.27	0.32	0.14	0.20	0.40	0.54	0.46	0.11	0.45	0.29	0.34	0.31	0.50	1.93	28.56	40.86	81.71	1.82	3.09	1.70
	(n = 4)	**Std**	0.03	0.04	0.04	0.03	0.04	0.06	0.04	0.05	0.07	0.06	0.05	0.02	0.02	0.03	0.21	2.01	2.28	26.64	0.10	0.19	0.11
	8	**Avg**	0.11	0.24	0.29	0.10	0.12	0.45	0.56	0.43	0.20	0.49	0.16	0.30	0.24	0.39	1.80	32.82	45.92	63.35	1.75	2.81	1.60
	(n = 4)	**Std**	0.01	0.04	0.06	0.02	0.02	0.03	0.03	0.08	0.08	0.04	0.05	0.04	0.01	0.02	0.33	3.68	6.03	6.96	0.05	0.39	0.20
	Avg/Site		0.12	0.26	0.32	0.13	0.17	0.45	0.56	0.46	0.15	0.48	0.20	0.32	0.28	0.43	1.91	31.88	43.66	70.34	1.73	2.97	1.72
	Std/Site		0.02	0.03	0.05	0.03	0.05	0.06	0.06	0.06	0.06	0.07	0.07	0.05	0.04	0.05	0.25	3.34	4.34	17.31	0.09	0.26	0.12

Table 2.4 (continued)

			Morphometric variables																			
	Shell		Values divided by H								Values divided by L				Angle measurement							
Site	forms		3	5	6	8	9	10	14	18	19	20	21	22	23	25	26	27	28	Vol	L/H	L/W	H/W
2	1	Avg	0.10	0.26	0.28	0.48	0.11	0.44	0.41	0.16	0.48	0.23	0.35	0.34	0.44	1.74	29.21	36.31	13.73	78.48	1.75	3.03	1.74
(n = 19)	(n = 5)	Std	0.03	0.03	0.03	0.72	0.02	0.05	0.03	0.02	0.09	0.06	0.03	0.05	0.05	0.14	1.60	2.48	6.56	15.00	0.05	0.19	0.15
	2	Avg	0.10	0.28	0.36	0.19	0.19	0.53	0.48	0.19	0.58	0.17	0.36	0.39	0.47	2.06	31.21	36.96	6.61	105.20	1.64	2.86	1.75
	(n = 5)	Std	0.03	0.12	0.05	0.04	0.04	0.05	0.06	0.04	0.09	0.06	0.03	0.06	0.06	0.27	2.83	2.43	7.36	25.37	0.03	0.12	0.06
	4	Avg	0.14	0.33	0.35	0.18	0.18	0.45	0.52	0.32	0.55	0.32	0.46	0.42	0.56	2.18	28.92	36.71	18.27	127.64	1.87	3.45	1.85
	(n = 3)	Std	0.03	0.04	0.04	0.02	0.02	0.04	0.09	0.11	0.07	0.08	0.07	0.07	0.08	0.34	2.96	0.57	17.48	12.17	0.13	0.10	0.09
	6	Avg	0.09	0.25	0.27	0.11	0.13	0.31	0.40	0.31	0.52	0.21	0.27	0.33	0.38	1.68	31.90	35.74	0.00	54.15	2.01	3.38	1.68
	(n = 1)	Std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	8	Avg	0.09	0.29	0.31	0.18	0.14	0.45	0.45	0.26	0.67	0.16	0.34	0.41	0.49	1.87	28.30	33.17	1.15	100.02	1.90	3.07	1.62
	(n = 5)	Std	0.02	0.02	0.06	0.06	0.03	0.07	0.07	0.05	0.12	0.04	0.05	0.07	0.05	0.28	2.76	2.35	2.57	22.77	0.06	0.20	0.11
	Avg/Site		0.10	0.28	0.32	0.26	0.15	0.46	0.46	0.23	0.57	0.21	0.36	0.38	0.47	1.92	29.59	35.69	8.54	97.66	1.79	3.08	1.72
	Std/Site		0.03	0.07	0.05	0.37	0.04	0.07	0.07	0.08	0.11	0.08	0.06	0.06	0.07	0.28	2.57	2.54	10.03	26.44	0.13	0.25	0.12
3	1	Avg	0.08	0.20	0.26	0.06	0.10	0.31	0.44	0.36	0.17	0.35	0.18	0.27	0.22	0.29	1.48	36.33	43.43	35.82	1.77	3.22	1.82
(n = 20)	(n = 7)	Std	0.02	0.05	0.05	0.01	0.02	0.10	0.11	0.07	0.06	0.11	0.07	0.06	0.05	0.07	0.28	3.63	3.38	16.51	0.10	0.30	0.11
	2	Avg	0.08	0.24	0.37	0.15	0.14	0.52	0.72	0.52	0.21	0.59	0.15	0.32	0.36	0.42	2.17	35.98	40.13	97.62	1.61	2.82	1.74
	(n = 6)	Std	0.02	0.05	0.08	0.04	0.03	0.11	0.15	0.09	0.04	0.14	0.04	0.05	0.05	0.08	0.38	3.70	4.60	39.26	0.05	0.29	0.17
	3	Avg	0.09	0.25	0.31	0.14	0.14	0.44	0.60	0.46	0.26	0.52	0.16	0.32	0.33	0.38	1.91	35.91	39.56	86.01	1.69	2.60	1.55
	(n = 6)	Std	0.02	0.04	0.05	0.04	0.06	0.07	0.11	0.07	0.06	0.11	0.05	0.07	0.06	0.08	0.29	4.42	4.44	31.47	0.05	0.13	0.11
	9	Avg	0.07	0.24	0.30	0.14	0.13	0.55	0.73	0.47	0.21	0.72	0.15	0.28	0.35	0.49	1.94	29.34	38.37	111.88	1.74	2.72	1.56
	(n = 1)	Std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Avg/Site		0.09	0.23	0.31	0.11	0.13	0.42	0.59	0.44	0.21	0.49	0.17	0.30	0.30	0.37	1.84	35.75	41.03	73.22	1.70	2.89	1.70
	Std/Site		0.02	0.05	0.07	0.05	0.04	0.13	0.17	0.10	0.06	0.16	0.06	0.06	0.08	0.10	0.41	3.90	4.22	39.83	0.09	0.35	0.17

Table 2.4 (continued)

			Morphometric variables																			
	Shell		Values divided by H							Values divided by L					Angle measurement							H/
Site	forms		3	5	6	8	9	10	14	18	19	20	21	22	23	25	26	27	28	Vol	L/H	L/W	W
4	1	Avg	0.06	0.16	0.22	0.08	0.09	0.30	0.42	0.34	0.13	0.35	0.15	0.21	0.21	0.26	1.41	36.78	43.53	29.66	1.75	2.98	1.70
(n = 20)	(n = 4)	Std	0.01	0.04	0.06	0.03	0.03	0.11	0.15	0.10	0.05	0.10	0.05	0.05	0.06	0.07	0.42	2.36	4.30	16.97	0.12	0.42	0.12
	2	Avg	0.06	0.13	0.28	0.13	0.13	0.36	0.47	0.40	0.13	0.36	0.15	0.25	0.23	0.28	1.66	39.66	45.22	38.55	1.56	2.57	1.64
	(n = 4)	Std	0.01	0.08	0.03	0.02	0.01	0.01	0.02	0.03	0.03	0.07	0.01	0.04	0.02	0.02	0.13	1.49	3.44	4.77	0.03	0.21	0.11
	3	Avg	0.11	0.29	0.29	0.15	0.17	0.47	0.61	0.45	0.24	0.54	0.20	0.36	0.36	0.41	1.92	32.88	36.88	86.50	1.76	2.98	1.69
	(n = 4)	Std	0.02	0.04	0.03	0.02	0.05	0.04	0.04	0.03	0.03	0.11	0.06	0.03	0.02	0.05	0.13	4.09	1.48	8.55	0.02	0.17	0.09
	5	Avg	0.06	0.16	0.19	0.10	0.07	0.31	0.43	0.29	0.18	0.45	0.15	0.23	0.26	0.33	1.20	27.05	33.04	25.88	2.06	3.83	1.86
	(n = 2)	Std	0.01	0.04	0.04	0.00	0.00	0.02	0.00	0.03	0.02	0.07	0.03	0.02	0.01	0.03	0.14	0.69	1.63	4.43	0.04	0.13	0.02
	7	Avg	0.08	0.26	0.30	0.20	0.16	0.51	0.64	0.45	0.10	0.48	0.19	0.31	0.34	0.41	1.88	33.19	38.17	77.11	1.64	2.66	1.63
	(n = 2)	Std	0.02	0.08	0.01	0.07	0.02	0.13	0.15	0.00	0.04	0.04	0.03	0.09	0.03	0.11	0.04	7.13	2.19	21.88	0.04	0.15	0.13
	8	Avg	0.06	0.17	0.21	0.10	0.10	0.31	0.41	0.33	0.10	0.39	0.18	0.23	0.24	0.32	1.37	31.31	38.95	37.03	1.79	2.80	1.56
	(n = 4)	Std	0.02	0.02	0.03	0.01	0.00	0.02	0.01	0.04	0.02	0.09	0.05	0.03	0.05	0.03	0.15	3.66	6.12	5.18	0.05	0.18	0.07
	Avg/Site		0.07	0.19	0.25	0.12	0.12	0.37	0.49	0.38	0.15	0.42	0.17	0.26	0.27	0.33	1.58	34.15	40.04	48.65	1.74	2.91	1.66
	Std/Site		0.02	0.07	0.05	0.04	0.04	0.10	0.12	0.08	0.06	0.10	0.04	0.07	0.07	0.08	0.32	4.90	5.25	25.66	0.15	0.41	0.12

Table 2.4 (continued)

			Morphometric variables																			
	Shell		Values divided by H								Values divided by L					Angle measurement				Vol	L/H	L/W	H/W
Site	forms		3	5	6	8	9	10	14	18	19	20	21	22	23	25	26	27	28				
5	1	Avg	0.05	0.13	0.27	0.06	0.16	0.29	0.41	0.40	0.09	0.31	0.21	0.23	0.24	0.26	1.65	42.38	43.74	34.00	1.75	2.81	1.61
(n = 20)	(n = 6)	Std	0.02	0.04	0.05	0.03	0.04	0.06	0.09	0.07	0.04	0.11	0.08	0.02	0.05	0.07	0.27	3.95	1.03	13.03	0.11	0.30	0.15
	2	Avg	0.09	0.23	0.33	0.12	0.18	0.46	0.62	0.49	0.28	0.40	0.17	0.31	0.31	0.35	2.05	39.97	43.38	60.20	1.52	2.71	1.79
	(n = 5)	Std	0.02	0.02	0.05	0.03	0.01	0.06	0.03	0.06	0.22	0.11	0.09	0.04	0.03	0.05	0.22	5.19	2.42	6.62	0.09	0.13	0.06
	3	Avg	0.10	0.24	0.32	0.12	0.23	0.42	0.53	0.46	0.08	0.36	0.18	0.33	0.26	0.38	1.93	35.74	46.30	54.62	1.61	2.74	1.70
	(n = 4)	Std	0.02	0.04	0.03	0.04	0.06	0.03	0.03	0.05	0.01	0.03	0.04	0.03	0.02	0.04	0.21	4.08	1.94	10.83	0.04	0.10	0.08
	4	Avg	0.10	0.24	0.26	0.12	0.21	0.40	0.50	0.40	0.13	0.37	0.22	0.33	0.30	0.38	1.64	31.57	37.95	50.48	1.82	3.17	1.74
	(n = 4)	Std	0.03	0.04	0.03	0.03	0.03	0.02	0.04	0.04	0.04	0.08	0.09	0.05	0.04	0.04	0.18	3.29	1.03	15.11	0.03	0.20	0.10
	7	Avg	0.08	0.18	0.38	0.08	0.21	0.29	0.37	0.48	0.07	0.19	0.30	0.32	0.21	0.32	1.99	41.46	53.41	42.95	1.62	2.68	1.66
	(n = 1)	Std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Avg/Site		0.08	0.20	0.30	0.10	0.19	0.38	0.50	0.44	0.14	0.35	0.20	0.30	0.27	0.34	1.82	38.24	43.49	48.42	1.67	2.84	1.70
	Std/Site		0.03	0.06	0.05	0.04	0.04	0.08	0.10	0.06	0.13	0.10	0.08	0.06	0.04	0.07	0.27	5.61	3.95	14.87	0.14	0.26	0.12

Table 2.4 (continued)

			Morphometric variables																			
	Shell		Values divided by H								Values divided by L					Angle measurement						
Site	forms		3	5	6	8	9	10	14	18	19	20	21	22	23	25	26	27	28	Vol	L/H	L/W	H/W
6	1	Avg	0.10	0.23	0.32	0.09	0.16	0.39	0.50	0.43	0.10	0.41	0.18	0.31	0.26	0.38	1.79	33.52	44.67	54.39	1.72	2.95	1.72
(n = 19)	(n = 5)	Std	0.02	0.04	0.07	0.03	0.06	0.09	0.11	0.08	0.03	0.13	0.04	0.05	0.06	0.08	0.32	2.22	4.75	25.79	0.07	0.12	0.12
	2	Avg	0.09	0.26	0.36	0.14	0.18	0.47	0.64	0.51	0.12	0.48	0.12	0.31	0.31	0.38	2.14	38.07	44.57	66.04	1.53	2.69	1.76
	(n = 4)	Std	0.04	0.06	0.02	0.04	0.04	0.09	0.08	0.04	0.07	0.05	0.02	0.07	0.05	0.04	0.16	1.59	5.73	14.73	0.03	0.20	0.14
	3	Avg	0.07	0.18	0.27	0.12	0.15	0.35	0.47	0.40	0.13	0.41	0.15	0.21	0.23	0.31	1.66	37.19	45.19	38.20	1.62	2.67	1.65
	(n = 4)	Std	0.02	0.03	0.03	0.02	0.02	0.05	0.07	0.04	0.03	0.05	0.05	0.03	0.03	0.04	0.16	5.40	5.27	5.60	0.03	0.14	0.10
	4	Avg	0.09	0.23	0.25	0.10	0.12	0.32	0.42	0.37	0.13	0.37	0.22	0.27	0.26	0.37	1.53	30.45	39.65	42.84	1.80	3.09	1.71
	(n = 4)	Std	0.01	0.02	0.00	0.02	0.01	0.04	0.04	0.02	0.06	0.09	0.04	0.02	0.02	0.02	0.09	0.44	2.37	4.11	0.06	0.17	0.12
	5	Avg	0.08	0.24	0.32	0.10	0.15	0.38	0.53	0.44	0.25	0.59	0.15	0.29	0.29	0.36	1.84	35.35	41.57	85.73	1.86	2.56	1.38
	(n = 2)	Std	0.01	0.02	0.08	0.06	0.03	0.05	0.02	0.08	0.07	0.00	0.03	0.01	0.02	0.05	0.35	1.21	3.65	27.88	0.00	0.51	0.27
	Avg/Site		0.09	0.23	0.30	0.11	0.15	0.38	0.51	0.43	0.13	0.43	0.17	0.28	0.27	0.36	1.79	34.80	43.38	54.30	1.69	2.83	1.68
	Std/Site		0.02	0.05	0.06	0.03	0.04	0.08	0.10	0.07	0.06	0.10	0.05	0.06	0.05	0.05	0.29	3.85	4.66	21.50	0.12	0.26	0.16
7	1	Avg	0.09	0.18	0.24	0.08	0.11	0.29	0.40	0.33	0.18	0.31	0.20	0.27	0.20	0.28	1.37	34.67	43.25	35.08	1.76	3.08	1.75
(n = 7)	(n = 6)	Std	0.02	0.05	0.08	0.03	0.04	0.08	0.10	0.11	0.05	0.12	0.03	0.07	0.04	0.08	0.44	2.23	4.94	21.92	0.06	0.28	0.11
	6	Avg	0.10	0.24	0.27	0.12	0.13	0.34	0.47	0.37	0.22	0.42	0.27	0.29	0.27	0.45	1.59	26.17	39.70	60.89	1.79	2.99	1.67
	(n = 1)	Std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Avg/Site		0.09	0.19	0.25	0.09	0.12	0.30	0.41	0.33	0.19	0.32	0.21	0.27	0.21	0.30	1.40	33.45	42.75	38.76	1.76	3.07	1.74
	Std/Site		0.02	0.05	0.07	0.03	0.04	0.08	0.09	0.10	0.05	0.11	0.04	0.06	0.05	0.10	0.41	3.80	4.70	22.26	0.06	0.26	0.11
8	1	Avg/ Site	0.13	0.35	0.44	0.16	0.22	0.58	0.76	0.58	0.27	0.65	0.38	0.41	0.41	0.56	2.42	31.43	40.08	162.24	1.78	3.79	2.13
(n = 5)		Std/ Site	0.02	0.041	0.04	0.02	0.05	0.02	0.04	0.03	0.07	0.19	0.17	0.03	0.04	0.06	0.11	2.64	2.88	31.14	0.04	0.23	0.12

Table 2.5. Results of factor analysis performed on morphometric variables that satisfied assumptions of multivariate statistical analyses. Variable names refer to the shell characters indicated in Table 2.2 that were divided by length and/or height measurements. Variables with asterisk were suggested to be better than a priori shell forms in distinguishing *E. complanata* from CFR.

Variable	Factor 1	Communality	Variable	Factor 2	Communality
* 14/H	0.816	1.000	* L/H	0.208	0.152
* 25	0.772	1.000	* 26	0.153	0.072
* 6/H	0.767	0.876	* logL/W	0.089	0.031
log9/H	0.601	0.497	* 27	0.069	0.074
logVol	0.597	0.752	log20/L	-0.05	0.055
10/H	0.549	0.679	H/W	-0.062	0.009
13/H	0.549	0.692	3/H	-0.369	0.197
23/L	0.479	0.526	log9/H	-0.369	0.497
22/L	0.443	0.518	log18/L	-0.392	0.153
19/L	0.379	0.397	log8/H	-0.484	0.324
21/L	0.368	0.381	21/L	-0.495	0.381
5/H	0.334	0.358	5/H	-0.497	0.358
log8/H	0.299	0.324	19/L	-0.504	0.397
27	0.263	0.074	6/H	-0.536	0.876
3/H	0.247	0.197	23/L	-0.544	0.526
log20/L	0.228	0.055	22/L	-0.567	0.518
26	0.22	0.072	14/H	-0.578	1.000
H/W	0.07	0.009	10/H	-0.614	0.679
log18/L	0.003	0.153	13/H	-0.626	0.692
logL/W	-0.153	0.031	logVol	-0.629	0.752
L/H	-0.33	0.152	25	-0.636	1.000

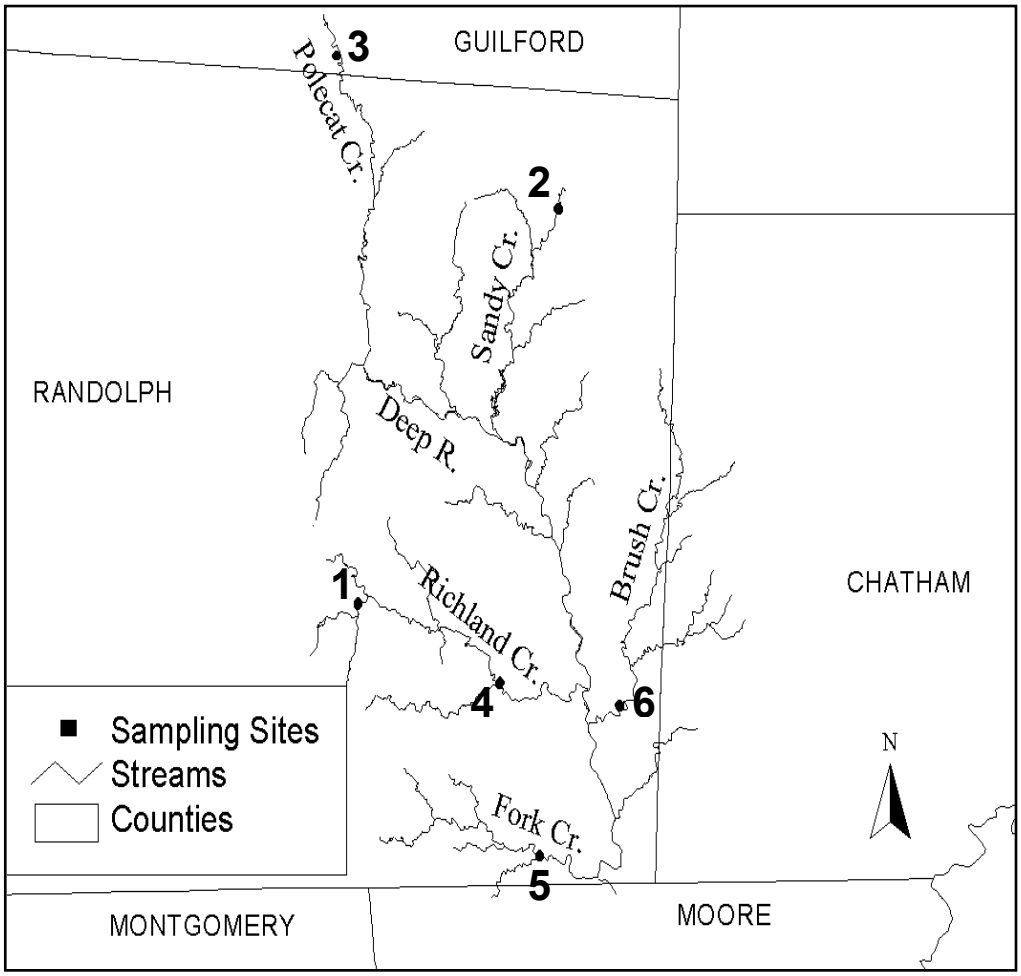


Figure 2.1. Sample site location.

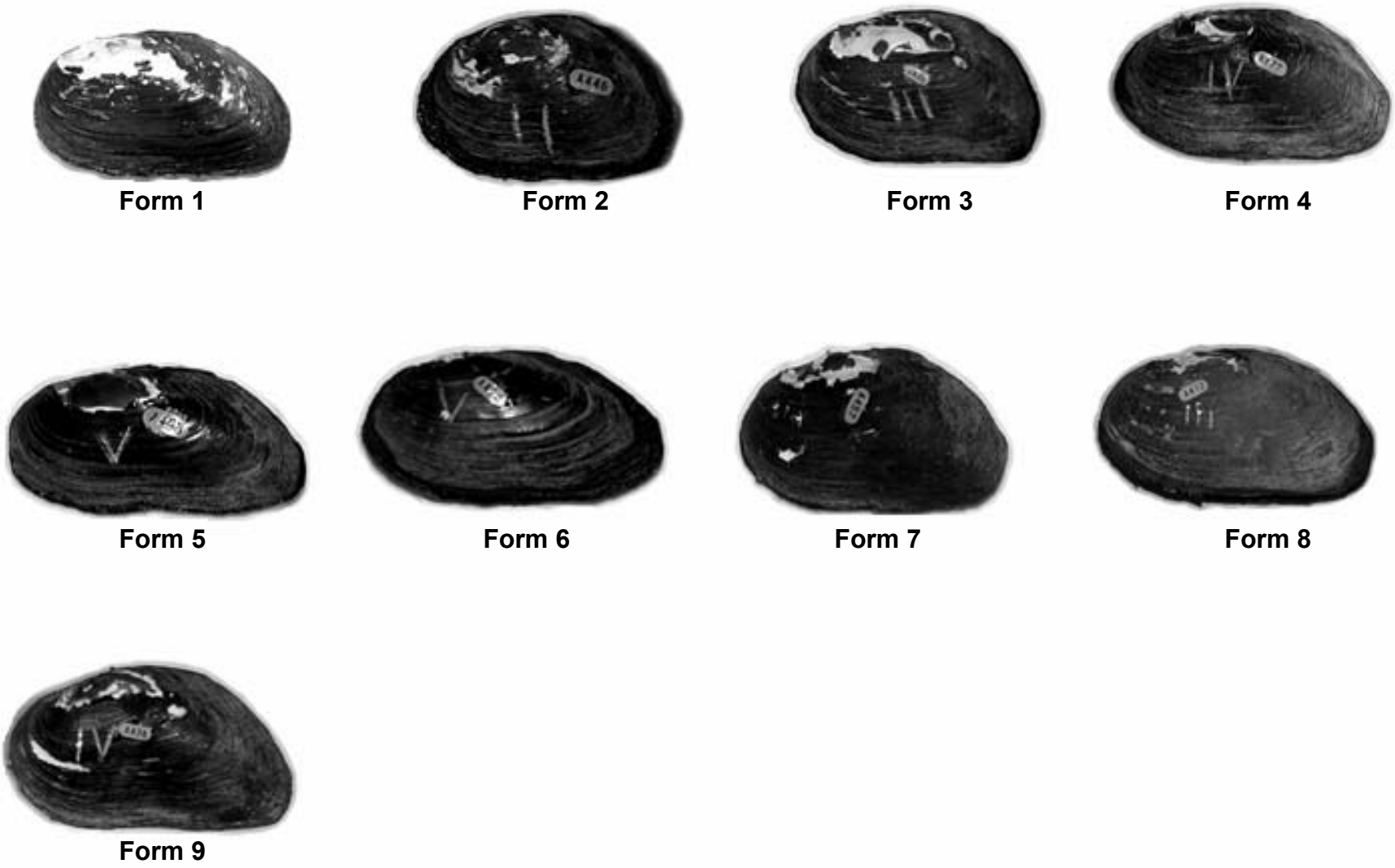
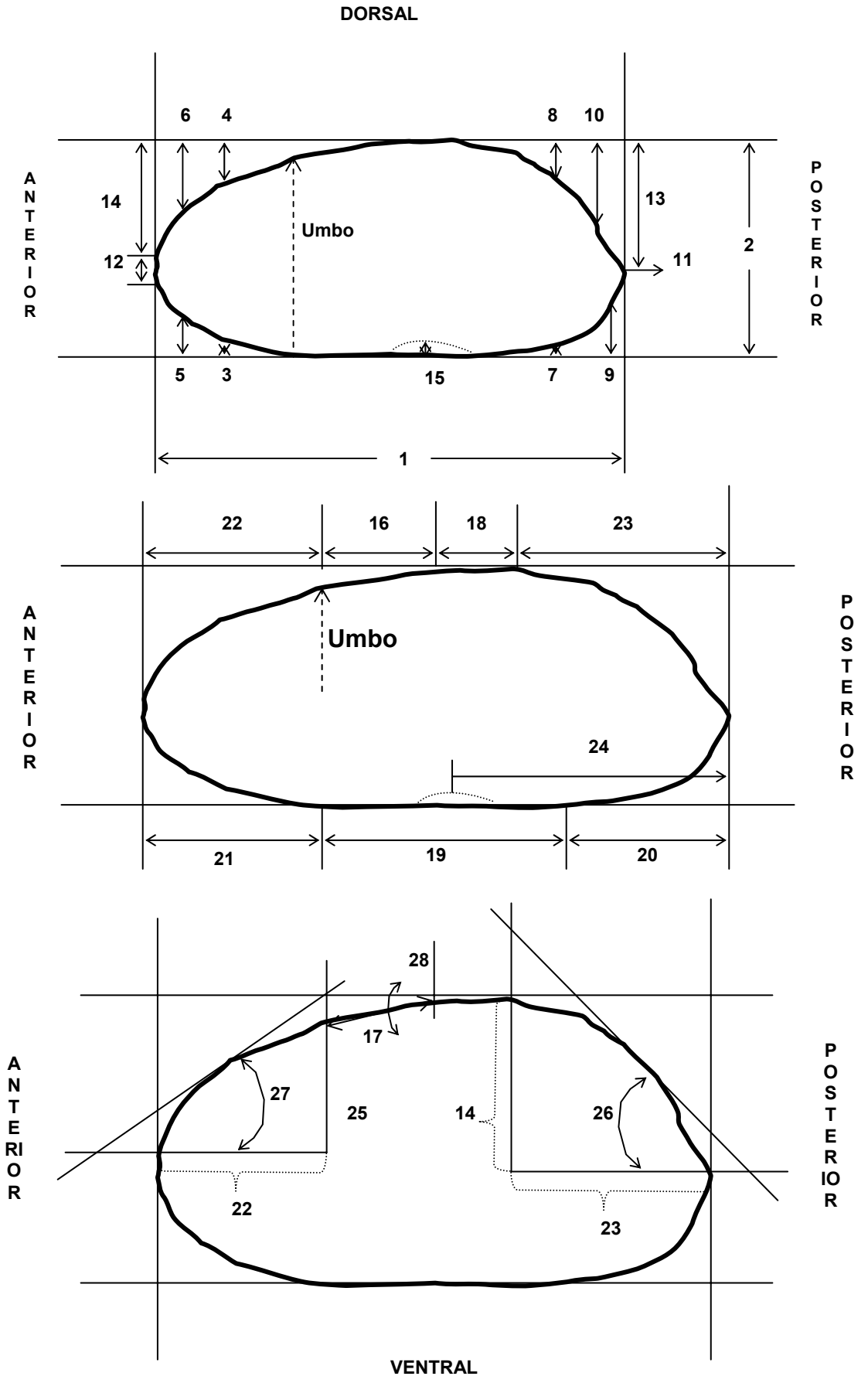


Figure 2.2. Representative pictures of the different *a priori* shell forms. Form 1 = 9.58 cm; Form 2 = 5.19 cm; Form 3 = 8.07 cm; Form 4 = 7.21 cm; Form 5 = 5.57 cm; Form 6 = 7.17 cm; Form 7 = 6.39 cm; Form 8 = 8.25 cm; Form 9 = 8.08 cm.

Figure 2.3. Shell outline enclosed by a rectangle and the different points that correspond to a specific shell character measurement based on Table 2.3. Numbers refer to shell characters listed in Table 2.3.


DORSAL
6
4
8
10
14
13
Umbo
12
11
2
ANTERIOR
POSTERIOR
5
3
15
7
9
1
22
16
18
23
Umbo
24
21
19
20
28
17
27
25
14
26
22
23
VENTRAL

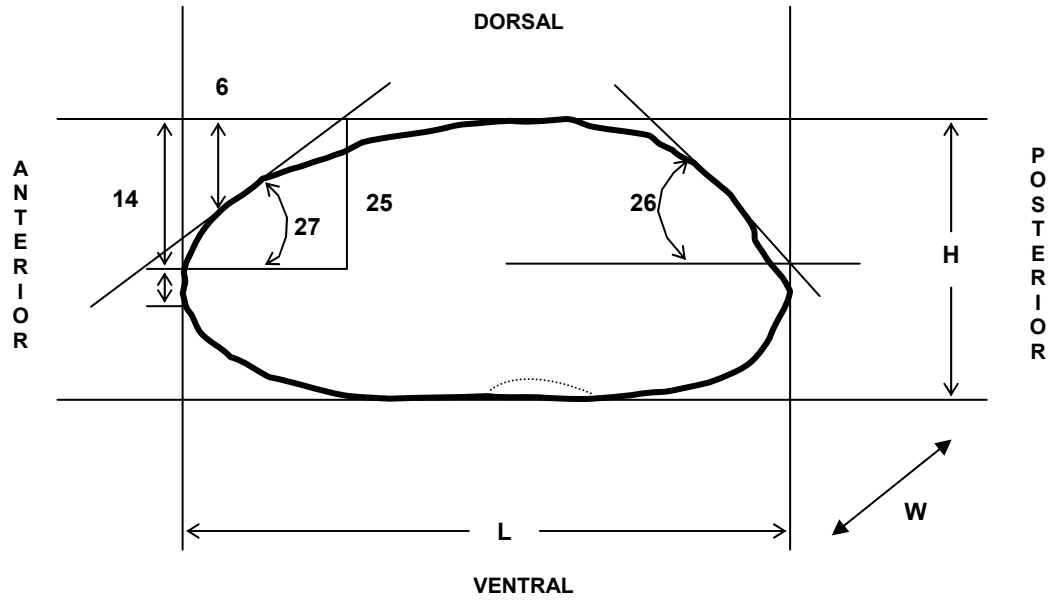


Figure 2.4. Orientation of the seven morphometric variables that had high loadings from factor analysis. Numbers correspond to shell characters described in Table 2.3.

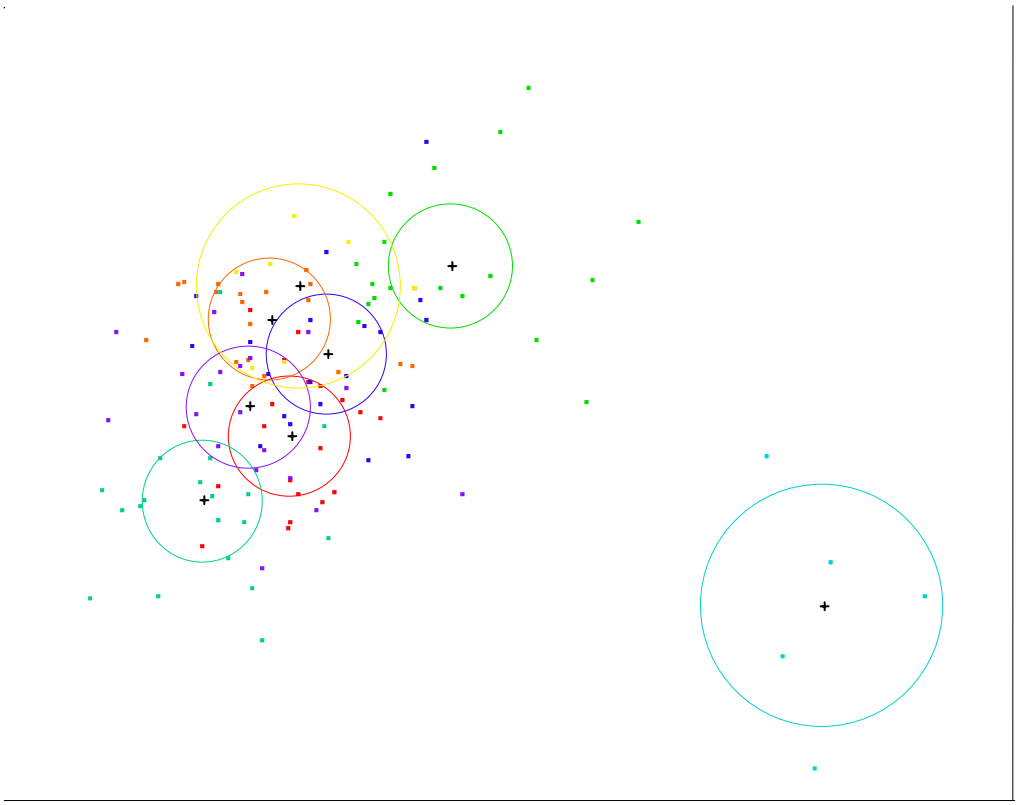


Figure 2.5. Discriminant plot based on site. Sites 1 to 7 are from Cape Fear River drainage; Site 8 is from Potomac River, Washington. Sample data points were shown as small dots. Circles represent 95% confidence interval for each site.

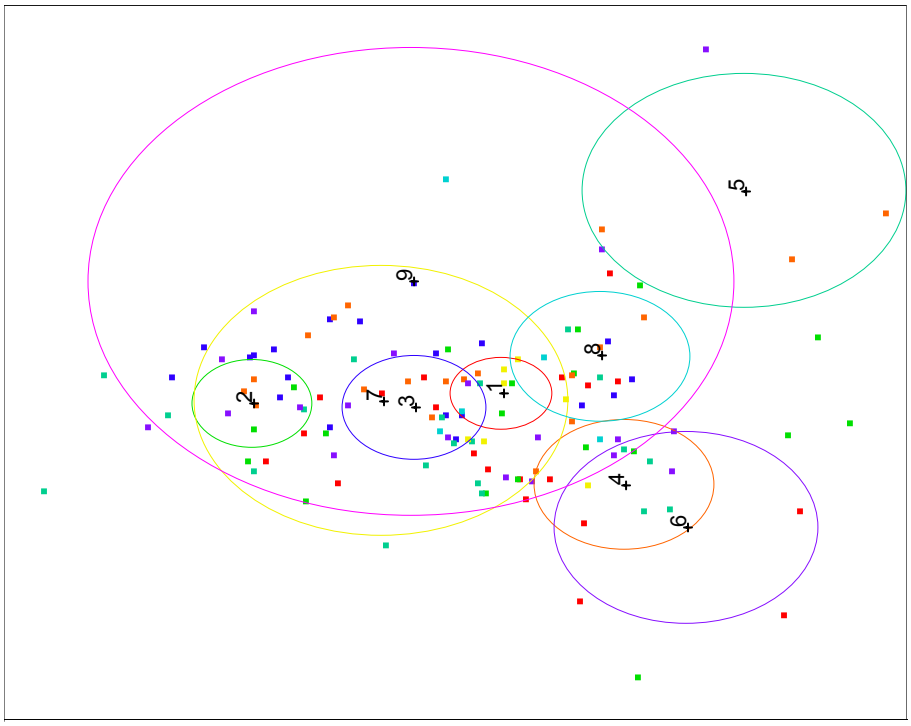
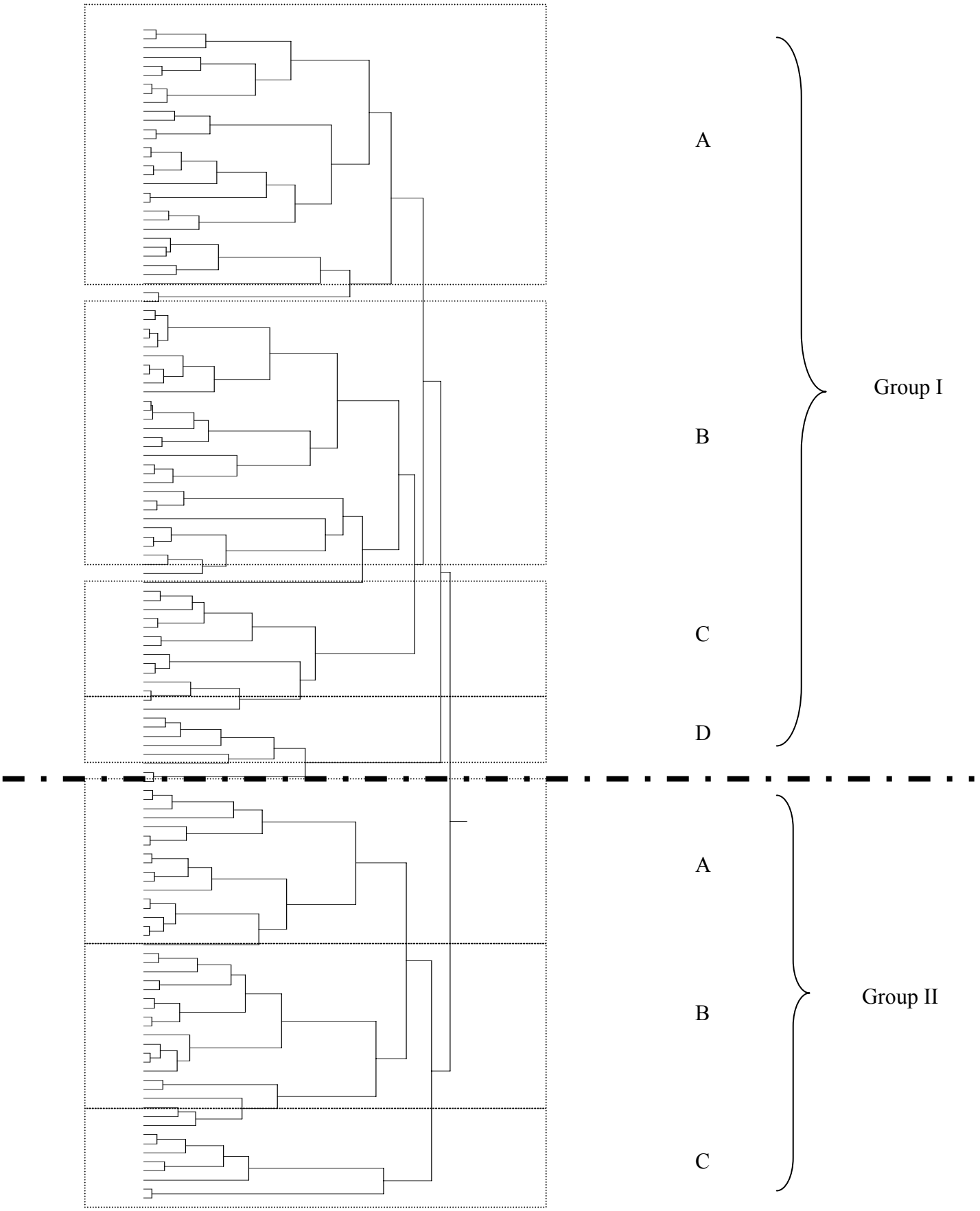


Figure 2.6. Discriminant plot according to shell morphology. Shell form numbers and description are based on Table 2.2. Sample data points were shown as small dots. Circles represent 95% confidence interval for each site.

Figure 2.7. Heirarchical cluster diagram of collected individuals from Cape Fear River drainage. Group I, sub-groups A to C were *E. complanata*-like; Group I, sub-group D, is *E. complanata*; Group II, sub-groups A to C, were non-*E. complanata*.


A
B
C
D
Group I
A
B
C
Group II

CHAPTER 3

PRELIMINARY EVIDENCE OF MORPHOLOGIC AND GENETIC VARIATION IN THE FRESHWATER MUSSEL, *ELLIPTIO COMPLANATA* (LIGHTFOOT 1786) FROM THE CAPE FEAR RIVER, NCreliminary evidence of morphologic and genetic variation in the freshwater mussel, *Elliptio complanata* (Lightfoot 1786) from the Cape Fear River, NC

R.A. Molina[1], M.E. Raley [1,2], A.E. Bogan[2] and J.F. Levine[1]

[1]Population Health and Pathobiology department, College of Veterinary Medicine, North Carolina State University, 4700 Hillsborough St., Raleigh NC 27606, Tel. #: (919) 513-6302. Email: rmolina@unity.ncsu.edu, revmolina@yahoo.com, Jay_Levine@ncsu.edu

[2]North Carolina Museum of Natural Sciences, Research Laboratory, 4301 Reedy Creek Road, Raleigh, NC 27607, Tel. #: (919) 733-7450 ext 753. Email: morgan.raley@ncmail.net, arthur.bogan@ncmail.net

ABSTRACT

Seven morphometric variables and cytochrome oxidase I (COI) sequence demonstrated variation within the freshwater mussel *Elliptio complanata*. Discriminant analysis and genetic diversity of the COI sequence revealed morphometric and genetic differences between samples from the Cape Fear River (CFR) and a topotype site (Potomac River, Washington D.C.). Sequence diversity supported the earlier hypothesis regarding geographical delineation of *E. complanata* (Matteson 1948). In addition, molecular diversity indices suggested directional gene flow in structuring genetic variation, similar to that found in some freshwater fish populations. Although cluster analysis of shell morphometry and genetic sequence data seem to support the existence of four different groups within the sampled CFR populations, there was no correlation found between the two datasets. The

lack of correlation between morphometry and genetic data was explained in the light of *E. complanata* shell plasticity, making identification based on the seven shell variables unreliable and confirming that genetic data is more appropriate in distinguishing the highly plastic *E. complanata*.

INTRODUCTION

Of the close to 300 freshwater mussel species in North America (Turgeon et al. 1998), the majority are found in the southeast U.S (Turgeon et al. 1998). The most common among these is the complex of species in the *Elliptio complanata* group. The diversity of shell forms has led earlier workers to assign subspecies and ecophenotypes (Matteson 1948) and/or synonymous species (Johnson 1970). To add to this confusion, Davis and coworkers (1981) suggested the possibility that there maybe hybridization among the many distinct species of the *E. complanata* group, based on the occurrence of sympatric phenotypes found in streams and lakes in the southeast U.S. (particularly in North Carolina, South Carolina, Georgia and Florida). Although a typical description of *E. complanata* had already been described and recognized, apparent shell intergradations are common in the southeast U.S. This variation has consistently made identification of species in this complex somewhat difficult. This apparent morphologic plasticity is also evident when attempting to identify other species within the genus *Ellipito*. Bogan and co-workers (2003) reported that the entire genus of *Elliptio* in the southern Atlantic region is not monophyletic. They suggested that based on COI, NDI and combined COI and NDI sequence data that the previously recognized taxa *E. lanceolata* and *E. steinstansana* were not members of the genus. *Elliptio lanceolata* appears to be more closely related with the genus *Lampsilis*. *Elliptio*

steinstansana appears to be more closely related to *Pleurobema collina*. In addition, the species *E. spinosa* was appropriately placed in *Elliptio* and the placement of *E. icterina* remained undetermined (Bogan et al. 2003). Bogan and co-workers (2003) add that *E. complanata* was more appropriately grouped with *E. roanokensis*, *U. quadridentatus*, *U. ratus* and *U. hepaticus*, which are synonymous species of *E. complanata* (for *U. ratus*) and *E. icterina* (for *U. hepaticus*). These studies highlighted the difficulty differentiating *E. complanata* forms from other *Elliptio* species. In an effort to reduce the confusion associated with the identification of *E complanata*, we examined the phenotypic and genotypic variation of individuals collected in a single river basin, the Cape Fear River Basin (NC). Individuals identified as *E. complanata* based on general morphology were further characterized using seven morphometric criteria and mitochondrial DNA analysis.

METHODOLOGY

Sample location and collection

The cohort of *E. complanata* examined was comprised of all animals collected in the Cape Fear basin utilized for a previously performed morphometric analysis (Molina et al., chapter two). From this group of animals, only *E. complanata* for which amplified DNA sequences were obtained were chosen for further morphometric and mitochondrial DNA analyses (Table 3.1).

Elliptio complanata individuals were collected from a total of eight sites. Seven sites were located near bridge crossings of the Cape Fear River (CFR) drainage as previously described (Molina et al., chapter one). The 300-m reaches immediately upstream and

downstream of each CFR road crossing were considered one site (Figure 3.1). Sampling within CFR was performed on sites that were located on various creeks along in the drainage except for one where two sites were chosen along one creek (Table 3.1). Six of the seven CFR sites were located southwest of Raleigh, NC (sites one to six) while one was located to the westward (site seven). Cape Fear River specimens were compared against topotypic specimens of *E. complanata* from the Potomac River near Washington, D.C. (courtesy of the North Carolina Museum of Natural Sciences).

Morphometric measurement and analyses

Seven morphometric variables following Molina et al. (chapter one) were taken from individual photographs of *E. complanata* samples, using the measure tool of ArcView GIS Version 3.2a. Molina and co-workers (chapter one) suggested that these measurements might be able to distinguish different *E. complanata* populations. The seven landmarks describe mussel shell measurements based on posterior and anterior angles and shell obesity (length/height and log-transformed length/width) (Figure 3.2). Variables were transformed, as necessary, to satisfy assumptions of multivariate normality and pairwise linearity following Johnson and Wichern (2002).

Multivariate analyses (discriminant and cluster analyses) were performed on the seven morphometric measurements to summarize and describe mussel morphometry. Discriminant analysis is a multivariate statistical tool that separates a dataset based on discriminants (Johnson and Wichern 2002). Specifically, discriminant analysis was used to separate *E. complanata* individuals according to site. The 'Discriminant' option of the

'Multivariate Method', under the 'Analyze' option in SAS-JMP (release 5.0) was utilized to perform this analysis.

Heirarchical cluster analysis based on Ward's distance method was utilized to determine the natural groupings created by the dataset for comparison with maximum parsimony results of sequence data. This analysis was performed using the 'Cluster' command, under the 'Multivariate' option of the program MINITAB Release 12.23.

Mitochondrial sequencing and data analysis

Mussels were euthanized and placed in 100% ethanol until tissue collection could be performed the following day. Shells were then forced open with a blunt knife and adductor muscles were cut to allow for easy access to the soft anatomy. Snips of foot, mantle and adductor tissues were removed for molecular analysis. Twenty-five mg of ethanol-preserved mantle, adductor and foot tissues were used for DNA extraction, which was performed using the DNeasy extraction kit (QIAGEN, Germantown, Maryland). Extracted DNA was run on 1% agarose gels stained with ethidium bromide.

We sequenced 609 bp of a 710 bp fragment of the cytochrome oxidase I (COI) mitochondrial gene. The forward and reverse primers reported by Folmer et al. (1994) were used in amplifying and sequencing *E. complanata* COI region. Initial sequencing attempts were first conducted on a few individuals to optimize the PCR steps before sequencing the rest of the samples.

Polymerase chain reaction (PCR) amplification reactions containing 2.5 μl each of 10 μM LC01490 (forward primer) and 10 □M HC02198 (reverse primer), 20 μl Eppendorf® Master Mix *Taq* polymerase, 23 μl sterile distilled H_2O and 2 μl DNA template in a total

volume of 50 μl were run for 40 cycles with an annealing temperature of 48^{o} to 50^{o}C. Product quality for PCR analysis was assessed on 1% ethidium bromide stained agarose gel. Samples with a single, well-defined band of approximately ~710 bp were cleaned using QIAquick® PCR Purification Kit (QIAGEN, Germantown, Maryland). Sequencing reactions were performed on purified products using the BigDye® Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, California) with LC01490 and HC02198 primers. Products were cleaned using DyeEx® spin column (QIAGEN, Germantown, Maryland) and visualized on an ABI377 sequencer (Applied Biosystems Inc., Foster City, California). Sequences were then aligned and edited using Sequencher 4.1 (Gene Codes Corp., Michigan). Proofread sequences were utilized for (1) genetic diversity analysis, using the freeware program, Arlequin ver. 2.000 (Schneider et al. 2000), and (2) maximum parsimony and bootstrap analyses (Phylogenetic Analysis Using Parsimony (PAUP*4.0) software (Swofford 2000). Sequences from 62 individuals were transformed into 25 haplotypes using the Collapse v.1.1 (http://inbio.byu.edu/Faculty/kac/crandall_lab/programs.htm). Haplotype creation made it easy for phylogenetic trees to be created and read.

Proofread sequences were grouped according to site and creek to estimate genetic diversity by geographical location. Molecular diversity indices, π, and nucleotide diversity were calculated at these levels. Indices based on mean number of pairwise differences (π) between all pairs of haplotypes, and average gene diversity over loci (nucleotide diversity), were measures used to quantify genetic diversity. These measures allow for comparison between all pairs of haplotypes found in the sample (in the case of π) and over all loci (in the case of nucleotide diversity) (Schneider et al., 2000). Analysis of molecular variance

(AMOVA) and population pairwise genetic distances (pairwise F_{ST}) were also performed at the two levels to differentiate the populations (Schneider et al., 2000).

Correlation between morphometry and DNA sequence

Correlation between morphometry and genetic groupings was performed through the least squares regression method. Morphometric distance based on Ward's distance method and pairwise similarity of COI sequence data were data used for correlation analysis (Husseneder and Grace 2001). Arc-sine and log-transformations of variables were performed to satisfy assumptions of normality and homoscedasticity.

RESULTS

Morphometric analysis

The collected *E. complanata* individuals showed shell variation according to measures of obesity, and posterior and anterior angles (Table 3.2). Sites four and eight had the lowest and highest value for variable six (dorsal anterior tangent). Sites four and six had the lowest measurement and site eight the highest record for variables 14 and 25 (dorsal departure from anterior vertical congruence and vertical distance from anterior vertical contact, respectively). For variable 26 (posterior angle), site eight was the lowest and the rest of the sites had values ranging from 32^{o} to 40^{o} (site six). Angle 27 (anterior angle) showed the largest and smallest anterior angle measures for sites one and two, respectively. Mussels from site eight had the most swollen shells while those from site three had the most compressed form based on measures of shell obesity (*i.e.* length/height and log-transformed length/width).

Heirarchical cluster analysis of morphometry data showed four groups belonging to two major clusters (Figure 3.3). Cluster I had two sub-groups. Group A had the highest number of individuals in a group (n = 26), while group B had six individuals, five of which were the topotypes. Cluster II had two sub-groups. Group C had 21 individuals. Group D only had six individuals. Samples were then grouped according to cluster groups and morphometric variables were summarized and inspected for possible patterns that might explain the groupings. There seems to be distinct differences between the different cluster groups based on variable six (Figure 3.4). Values of samples from group D fell at the left end of the spectrum while those from group B occupied the opposite end. Samples from groups A and C occupied the middle section of the range. This same trend was apparent for variables 14 and 25. There seems to be indistinguishable differences among all groups in the remaining variables (L/H, logL/W, 26 and 27) since each group's value range overlapped with another.

There was a significant difference between all CFR sites (sites one to seven) vs. site eight when examined by discriminant analysis based on the different sites (Figure 3.5).

Mitochondrial DNA analysis

General statistics

Mitochondrial sequences (n = 62) were obtained from LCOI490 and HCO2198 (Folmer et al. 1994), which yielded 609 base pairs of unambiguous sequence. A total of 25 unique COI haplotypes were recovered (sequences will be submitted to Genbank upon acceptance of article) (Table 3.3). Nucleotide composition was dominated by T (41% ± 0.08) and G (28% ± 0.12), followed by A and C at 16% (± 0.09) and 15% (± 0.05),

respectively. There were a total of fifty-seven variable characters, twenty-two of which were informative for parsimony analysis. Sequence divergence ranged from a minimum of 0.16% to a maximum of 5.42% (Table 3.4). Maximum divergence occurred only once between haplotypes 496 vs. 456.

Phylogenetic estimates

There were three main branches emanating from the center of the unrooted maximum parsimony-based dendrogram (Figure 3.6). Each main branch had further subdivisions within, but because of their low bootstrap scores compared to the main branches, only the main branches that had high bootstrap scores were considered. The main branches formed three groups that were well supported by bootstrap analysis. Group C had the highest bootstrap score (97), followed by group A at 74 and group B at 58. Group A had 38 individuals, including the five topotypic specimens, and Groups B and C each had ten individuals. Since samples from group A contained the topotypes, these maybe considered the typical *E. complanata*. Groups D and E which, had one individual each, did not have any support thus, did not show any bootstrap score.

Population subdivision

Pairwise differences (π) among the sites showed differentiation between site eight and the rest of the CFR sites (Table 3.5). Significant differences between the upstream site (site three = Polecat Creek) and the downstream sites (sites one, four, five and six) were also apparent based on π. Within the CFR sites, sites three and four had the lowest and highest mean diversity over all loci (nucleotide diversity), respectively (Figure 3.7).

Partitioning among different sites yielded significant results with F_{ST} at 0.11 (p-value = 0.01 ± 0.003 from 1023 permutations). Analysis of molecular variance (AMOVA) at the

creek-level (within CFR) yielded F_{ST} = 0.063, with p-value (0.053 ± 0.007) approaching significance (1023 permutations).

Correlation of morphometric distances and COI sequence similarities

There is negative correlation between morphometry and sequence data (Figure 3.8). Negative correlation coefficient (r = - 0.103, p = 0.443) resulted from the bivariate fit of log-transformed morphometric distance and arc-sine transformed genetic similarity.

DISCUSSION AND CONCLUSION

Morphometry

Based on discriminant analysis of morphometry data obtained from the topotype and all the CFR sites, the *E. complanata* from CFR were different from what was historically described as *E. complanata*. This is consistent with earlier work (Matteson 1948) that suggests geographical differences in *E. complanata* found north and south of Washington. Animals found north of Washington appear to be monomorphic, while those found south of Washington are not.

Groups generated from cluster analysis of morphometric data suggest the existence of four major groups. However, when compared with dendrogram of sequence data, individual samples found in morphometry clusters did not match those found in COI groups. Non-significant correlation between the two analyses also corroborated the observation that individuals from both groupings do not match. Lack of correlation between the two analyses highlighted the difference between the nature of morphometry and sequence data. Morphometry-derived groups, based on the seven morphological characters used in the study

reflected clustering based on the mussel's plastic response to its specific habitat. Freshwater mussels, like most mollusc species, are identified through their shell characteristics but such features are usually difficult to discern. Within-species differences in shell characteristics are apparent (Ortmann 1920, Renard et al. 2000) because of high plasticity (Mitton 1977, Morton 1987) and response to the environment (Belanger 1991, Stites et al. 1995). Similarities in shell morphology are also noticeable among different species found in the same area. Occurrences such as these are often referred to as 'site effects' (Stiven and Alderman 1993). These cases are especially applicable to *E. complanata*, which seems to be morphologically plastic (Molina et al., chapter one). Because of this, *E. complanta* shell identification based solely on the shell characters utilized in this study, does not seem to be reliable. Groupings based on genetic data, on the other hand, may be considered as another assemblage that reflects inherent genetic characters of the individuals and as such, may be considered more appropriate than the seven shell landmarks used, for *E. complanata* because it is devoid of bias considering the plastic tendency of this species. The appropriateness of genetic data over morphology-based phylogeny was also demonstrated in the case of the fairy shrimp, *Chirocephalus diaphanus,* which was earlier considered monophyletic based on appendage morphology but was later reported as displaying genetic differentiation in contradiction with appendage morphology but consistent with heterogeneity observed in a new morphological character (Ketmaier et al. 2003).

COI Sequence

Sequence data supported site differences among samples between CFR and topotypes (see above). In terms of genetic diversity within a site, site eight (site where topotypes were

from) showed lack of genetic diversity among individuals, indicating sequence similarity among all the topotypes. Compared to the CFR sites, site eight showed genetic differentiation based on population pairwise F_{ST}. This corroborated the significant results of AMOVA at the site-level. These results suggest that based on COI sequence, topotypic specimens belong to a single species group different from the *E. complanata* collected from the CFR. These results seem to support the geographical pattern of *E. complanata* shell morphology – those found north of Washington seem to possess a single shell form while those found south of it, have more diverse morphology (Matteson 1948).

Molecular diversity indices point to genetic differences between upstream and downstream sites. Standard measures of genetic diversity suggested higher variability among downstream sites. Site four, which had the highest gene diversity, is located downstream of site one. The next two sites that had high diversity scores are sites five and six. The former was located downstream of Fork Creek and the latter was found downstream of Brush Creek. Mean pairwise difference between all pairs of haplotypes showed that sites downstream (*i.e.* two sites at Richland creek, Brush Creek and Fork Creek) was significantly different from Polecat creek, the northern-most site in the sampling area. Site differences such as these demonstrated the importance of directional gene flow (from upstream to downstream sites) in structuring genetic variation among populations. Whitehead and co-workers (2003) demonstrated that physical geographic factors such as direction of river flow are responsible for structuring genetic variation among populations of the freshwater fish, *Catostomus occidentalis* in California. By correlating high diversity values found in the downstream vs. upstream sites, the authors showed that biogeographical patterns were more responsible for structuring genetic variation within the California watershed, than contaminant exposure

history (Whitehead et al. 2003). If this trend is generally true for freshwater fishes then it could follow that interconnected watersheds are also responsible for structuring of genetic variation in freshwater mussels due to their parasitic mode of reproduction on fishes.

Site difference of genetic diversity seems to partially support the theory on shape and station suggested by Ortmann (1920). The theory suggested that certain species display morphological differences distinguished by shell obesity, depending on the type/location of water system they inhabit: more swollen shell form was found in larger rivers; more compressed forms were found in headwaters and smaller creeks, and shell intergrades between the forms were present in medium-sized water systems (Ortmann 1920). Swollen forms were also found in downstream areas and more compressed forms, in upstream location. This same pattern was apparent only in sampled sites in the CFR that were located along the same creek (*i.e.* Richland creek). In terms of measures of obesity, site four, located downstream of site one, had the highest shell diameter (length/height and log-transformed length/width) and the highest genetic diversity values. The remainder of the CFR sites does not seem to follow the pattern suggested by the theory. These CFR sites were located in smaller creeks. Additional sampling in the Deep River portion of CFR would facilitate further consideration of Ortmann's theory.

The number of animals available for this study was limited by state permitting guidelines and source and quality of DNA that allowed for successful amplification of the COI gene. As a result, fewer mussels were found in downstream sites. Despite this, the calculated estimates of genetic diversity do not suffer from inaccuracy because downstream populations had many polymorphic loci (Krauss 2000). However, statistical comparisons would be more robust if estimates were derived from a larger sample of individuals. It is

suggested that future studies involving genetic diversity estimates for aquatic species in general, be conducted on at least ten individuals to provide a lower probability of Type II error (or higher discriminatory power of the test) (Berg and Berg 2000). Thus, based on small sample size used in this study, estimates of genetic diversity were only suggestive of geographic directionality and still warrant further investigation.

Groups derived from mitochondrial analysis suggest the existence of four major groups within the sampled *E. complanata* from CFR. Only one of the four groups may be considered typical *E. complanata* based on its similarity with topotypic specimens. The other three groups need to be re-evaluated by molecular comparison with other *Elliptio* species to better define their systematic affinities. This finding suggests that the collected *E. complanata* specimens from CFR are not a cohesive group as suggested by earlier researchers. As reported by others, *E. complanata* may not be a monophyletic group (Davis et al. 1981, Bogan et al. 2003). Davis et al. (1981) recognized high interpopulation phenotypic diversity in their collections and suggested that this species group was radiating and probably undergoing hybridization among the different ecophenotypes based on high genetic similarity values (0.99 to 0.99), high polymorphism, and considerable heterozygosity values derived from allozyme electrophoresis. In a study of the genus *Elliptio*, Bogan and co-workers (2003) reported that the species *E. complanata* is not cohesive based on genetic sequences of two molecular data sets, COI and NDI.

The lack of cohesion within *E. complanata* that we have described suggests the need for further study of this species, and a thorough reevaluation of the genus (Bogan et al., 2003). Based on morphometry and genetic sequencing the *Elliptio complanata*-like species we collected in the CFR do not resemble topotype *E. complanata*. Additional work is needed

to expand and further characterize phenotypic characters. Internal anatomy such as stomach features has been utilized to differentiate *E. complanata* from other species and locales (Davis and Fuller 1981, Kat 1983b). Similarly genetic analysis of freshwater mussels needs to be broadened to include an examination of additional genes. Bogan and co-workers (2003) used short sections from the first subunit of NADH dehydrogenase (NDI) and cytochrome oxidase c (COI) genes of the mitochondrial region to distinguish several individuals from the genus *Elliptio*. These additional characters need to be incorporated with the techniques used in this study for future analysis of *E. complanata*.

ACKNOWLEDGEMENT

The authors would like to thank the following for their invaluable assistance: Chris Eads and all the biologists of the Freshwater Mussel Research team at the College of Veterinary Medicine (NCSU), Dr. Peter Bloomfield (NCSU-Statistics), Judith Johnson-Ratcliff (NCWRC) and Jonathan Molina (MPO-RSMAS/UM). Funding for this research was provided by NCDOT and USFWS. This work is dedicated to J. Isaac.

REFERENCE

Belanger, S.E. 1991. The effect of dissolved oxygen, sediment and sewage treatment plant discharges upon growth, survival and density of Asiatic clams. Hydrobiologia 208: 113-126.

Berg, D. J.and P. H. Berg. 2000. Conservation genetics of freshwater mussels: comments on Mulvey et al. Conservation Biology 14(6): 1920-1923.

Bogan, A.E. 2002. Workbook And Key To The Freshwater Bivalves of North Carolina. North Carolina Museum of Natural Sciences, Raleigh, NC, 101 pp, 10 color plates.

Bogan, A.E., J.M. Serb and C. Lydeard. 2003. Determination of the systematic position and relationships of the *Elliptio* complex (Mollusca: Bivalvia: Unionidae) with distributions in North and South Carolina. A Final Report submitted to USGS (Contract 01HQAG0185).

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. Fuller and C. Hesterman. 1981. Molecular genetics and speciation in *Elliptio* and its relationships to other taxa of North American Unionidae (Bivalvia). Biological Journal of the Linnean Society 15: 131-150.

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. Mol. Mar. Biotechnol. 3: 294-299.

Hussender, C. and J. K. Grace. 2001. Evaluation of DNA fingerprinting, aggression tests, and morphometry as tools for colony delineation of the Formosan Subterranean termite. Journal of Insect Behavior 14(2): 173-186.

Johnson, R.A. and D.W. Wichern. 2002. Applied Multivariate Statistical Analysis, fifth edition. Prentice Hall, New York, NY. 767 p.

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

Ketmaier, V., P. Zarattini, E. De Matthaeis, M. Cobolli and G. Mura. 2003. Intra- and inter-specific relationships in the six Italian species of the fairy shrimp genus *Chirocephalus*: combining allozyme and mtDNA data. J. Zool. Syst. Evol. Research 41: 276-285.

Krauss, S.L. 2000. Accurate gene diversity estimates from amplified fragment length polymorphism (AFLP) markers. Molecular Ecology 9: 1241-1245.

Matteson, M.R. 1948. The taxonomic and distributional history of the freshwater mussel *Elliptio complanatus* (Dillwyn, 1817). The Nautilus 61(4): 127-132.

Mitton, J.B. 1977. Shell colour and pattern variation in *Mytilus edulis* and its adaptive significance. Chesapeake Science 18: 387-390.

Morton, B.S. 1987. Polymorphism in *Corbicula fluminea* (Bivalvia: Corbiculoidae) from Hong Kong. Malacological Review 20: 105-127.

Ortmann, A.E. 1920. Correlation of shape and station in freshwater mussels (Naiades). Proc. Am. Phil. Soc. Philr. 269-312.

Renard,E., V. Bachmann, M.L. Cariou and J.C. Moreteau. 2000. Morphological and molecular differentiation of invasive freshwater species of the genus *Corbicula (Bivalvia, Corbiculidea)* suggest the presence of three taxa in French rivers.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Stites, D.L., A.C. Benke, D.M. Gillespie. 1995. Population dymanics, growth and production of the asiatic clam, *Corbicula fluminea*, in a blackwater river. Canadian Journal of Fisheries and Aquatic Sciences 52: 425-437.

Stiven, A.E. and J. Alderman. 1993. Environmental and genetic approaches to understanding the status of freshwater mussels: the analysis of site effects and genetic similarity. Final report submitted to NC Nongame Small Grants Program, NC Wildlife Resources Commission.

Swofford, D.L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer, Sunderland, MA.

Turgeon, 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, Second Edition. American Fisheries Society Special Publication, 25: 526 p.

Whitehead, A., S. L. Anderson, K.M. Kuivila, J.L. Roach. 2003. Genetic variation among interconnected populations of Catastomus occidentalis: implications for distinguishing impacts of contaminants from biogeographical structuring. Molecular Ecology 12: 2817-2833.

Table 3.1. Geographic locations of sampled sites. Numbers of individuals used in analyses are in parenthesis below site numbers. Legend for dominant substrate: Clay = 0, Silt = 1, Sand = 2, Gravel = 3, Cobble = 4, Boulder = 5, Bedrock = 6.

Counties	Latitude	Longitude	Stream	Site number	Site name	Dominant substrate	Specimen number
Randolph	35.65	-79.78	Richland Creek (upstream)	1 (n = 14)	Richland1	4.3	379 to 398
Randolph	35.79	-79.67	Sandy Creek	2 (n = 11)	Sandy	4.4	399 to 418
Guilford	35.92	-79.80	Polecat Creek	3 (n = 9)	Polecat	3.6	419 to 438
Randolph	35.61	-79.68	Richland Creek (downstream)	4 (n = 3)	Richland2	4.5	439 to 458
Randolph	35.53	-79.64	Fork Creek	5 (n = 7)	Fork	4.8	459 to 478
Randolph	35.60	-79.58	Brush Creek	6 (n = 7)	Brush	5.4	479 to 498
Alamance			Mary's Creek	7 (n = 6)	Mary	2	499 to 505
Washington, D.C.	39.00	77.25	Potomac River	8 (n = 5)	Topotype		506 to 510

Table 3.2. Summary statistics and description of the seven morphometric measurements grouped by site.

Morphometric Variables		Sites							
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
6 Dorsal Anterior Tangent	n	14	10	9	3	6	6	6	5
	Mean	0.32	0.32	0.33	0.22	0.27	0.31	0.23	0.44
	Std. Dev.	0.05	0.06	0.07	0.08	0.06	0.07	0.07	0.04
	Min	0.23	0.20	0.18	0.14	0.18	0.25	0.15	0.39
	Max	0.41	0.42	0.43	0.30	0.35	0.38	0.33	0.51
14 (Dorsal Departure from Anterior vertical Congruence)	n	14	10	9	3	6	6	6	5
	Mean	0.46	0.45	0.47	0.31	0.40	0.44	0.31	0.58
	Std. Dev.	0.06	0.07	0.11	0.13	0.08	0.08	0.08	0.03
	Min	0.39	0.34	0.27	0.19	0.27	0.36	0.20	0.55
	Max	0.56	0.55	0.58	0.45	0.51	0.53	0.41	0.62
25 (Vertical Distance from Anterior vertical Contact to 21*)	n	14	10	9	3	6	6	6	5
	Mean	1.92	1.90	1.94	1.31	1.65	1.86	1.31	2.42
	Std. Dev.	0.27	0.29	0.45	0.53	0.30	0.36	0.35	0.11
	Min	1.62	1.40	1.09	0.79	1.14	1.51	0.83	2.30
	Max	2.35	2.33	2.41	1.85	2.04	2.29	1.68	2.60
26 (Posterior angle)	n	14	10	9	3	6	6	6	5
	Mean	32.55	29.53	35.27	33.25	40.17	35.50	33.05	31.43
	Std. Dev.	3.36	3.02	3.95	5.38	7.14	3.15	4.00	2.64
	Min	26.92	25.23	28.88	27.54	27.26	31.04	26.17	27.45
	Max	38.73	34.20	40.96	38.23	49.11	40.06	37.12	34.46
27 (Anterior angle)	n	14	10	9	3	6	6	6	5
	Mean	44.06	36.03	39.98	37.83	42.97	42.37	41.63	40.08
	Std. Dev.	4.71	2.78	4.62	3.15	2.24	5.78	4.01	2.88
	Min	39.45	31.66	32.19	34.20	39.08	36.73	36.97	36.58
	Max	54.87	39.70	44.20	39.72	45.33	53.03	47.51	44.60
L/H (Length/Height)	n	14	10	9	3	6	6	6	5
	Mean	1.73	1.75	1.69	1.89	1.72	1.71	1.78	1.78
	Std. Dev.	0.10	0.13	0.11	0.22	0.12	0.14	0.05	0.04
	Min	1.58	1.61	1.56	1.66	1.56	1.53	1.74	1.73
	Max	1.96	1.99	1.91	2.09	1.88	1.86	1.86	1.82
LogL/W (Length/Width)	n	14	10	9	3	6	6	6	5
	Mean	0.47	0.48	0.44	0.52	0.47	0.45	0.49	0.58
	Std. Dev.	0.04	0.03	0.05	0.10	0.03	0.06	0.04	0.03
	Min	0.36	0.45	0.39	0.41	0.44	0.34	0.45	0.53
	Max	0.53	0.53	0.55	0.59	0.50	0.52	0.55	0.60

*21 = ventral anterior departure

Table 3.3. List of COI haplotypes and corresponding sample numbers.

Haplotype	Sample number
1	Ecomp379, Ecomp389, Ecomp400, Ecomp401, Ecomp402, Ecomp422, Ecomp425, Ecomp426, Ecomp428, Ecomp435, Ecomp436, Ecomp485
2	Ecomp381, Ecomp391, Ecomp474, Ecomp497, Ecomp505
3	Ecomp383
4	Ecomp386, Ecomp387, Ecomp392, Ecomp398, Ecomp442
5	Ecomp388, Ecomp397, Ecomp503, Ecomp506, Ecomp507, Ecomp508, Ecomp509, Ecomp510
6	Ecomp390, Ecomp479, Ecomp484, Ecomp500
7	Ecomp394
8	Ecomp395, Ecomp465
9	Ecomp399, Ecomp405, Ecomp427, Ecomp467
10	Ecomp407, Ecomp409
11	Ecomp408
12	Ecomp410
13	Ecomp412
14	Ecomp414
15	Ecomp432, Ecomp433
16	Ecomp456
17	Ecomp457
18	Ecomp459
19	Ecomp460, Ecomp461
20	Ecomp475
21	Ecomp495
22	Ecomp496
23	Ecomp498
24	Ecomp499
25	Ecomp501, Ecomp502

Table 3.4. Pairwise comparison of genetic distance for COI sequence data.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	**Ecomp379.**	-																														
2	**Ecomp381.**	0.02	-																													
3	**Ecomp383.**	0.02	0.00	-																												
4	**Ecomp386.**	0.00	0.02	0.02	-																											
5	**Ecomp387.**	0.00	0.02	0.02	0.00	-																										
6	**Ecomp388.**	0.00	0.02	0.02	0.00	0.00	-																									
7	**Ecomp389.**	0.00	0.02	0.02	0.00	0.00	0.00	-																								
8	**Ecomp390.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	-																							
9	**Ecomp391.**	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	-																						
10	**Ecomp392.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	-																					
11	**Ecomp394.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	-																				
12	**Ecomp395.**	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-																			
13	**Ecomp397.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	-																		
14	**Ecomp398.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	-																	
15	**Ecomp399.**	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	-																
16	**Ecomp400.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	-															
17	**Ecomp401.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	-														
18	**Ecomp402.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	-													
19	**Ecomp405.**	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	-												
20	**Ecomp407.**	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	-											
21	**Ecomp408.**	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	-										
22	**Ecomp409.**	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	-									
23	**Ecomp410.**	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	-								
24	**Ecomp412.**	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.00	-							
25	**Ecomp414.**	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.00	0.00	-						
26	**Ecomp422.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	-					
27	**Ecomp425.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	-				
28	**Ecomp426.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	-			
29	**Ecomp427.**	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.01	0.01	0.01	0.01	0.01	-		
30	**Ecomp428.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	-	
31	**Ecomp432.**	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	-

Table 3.4 (continued)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
32 Ecomp433.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
33 Ecomp435.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
34 Ecomp436.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
35 Ecomp442.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
36 Ecomp456.	0.04	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.04	0.05
37 Ecomp457.	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.01	0.01
38 Ecomp459.	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
39 Ecomp460.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
40 Ecomp461	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
41 Ecomp465.	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.01	0.01
42 Ecomp467.	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01
43 Ecomp474	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.02	0.02	0.02	0.01	0.02	0.02
44 Ecomp475.	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.02	0.02	0.02	0.01	0.02	0.02
45 Ecomp479.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
46 Ecomp484.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
47 Ecomp485.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
48 Ecomp495.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
49 Ecomp496.	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02
50 Ecomp497.	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.02	0.02	0.02	0.01	0.02	0.02
51 Ecomp498.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
52 Ecomp499.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
53 Ecomp500.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
54 Ecomp501.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
55 Ecomp502.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
56 Ecomp503.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
57 Ecomp505.	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.00	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.02	0.02	0.02	0.01	0.02	0.02
58 Ecomp506.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
59 Ecomp507.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
60 Ecomp508.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
61 Ecomp509.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
62 Ecomp510.	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00

Table 3.4 (continued)

		32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62
32	**Ecomp433.**	-																														
33	**Ecomp435.**	0.00	-																													
34	**Ecomp436.**	0.00	0.00	-																												
35	**Ecomp442.**	0.00	0.00	0.00	-																											
36	**Ecomp456.**	0.05	0.04	0.04	0.05	-																										
37	**Ecomp457.**	0.01	0.01	0.01	0.01	0.05	-																									
38	**Ecomp459.**	0.02	0.02	0.02	0.02	0.05	0.02	-																								
39	**Ecomp460.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	-																							
40	**Ecomp461**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	-																						
41	**Ecomp465.**	0.01	0.01	0.01	0.01	0.04	0.00	0.02	0.01	0.01	-																					
42	**Ecomp467.**	0.01	0.01	0.01	0.01	0.05	0.00	0.02	0.01	0.01	0.00	-																				
43	**Ecomp474**	0.02	0.02	0.02	0.02	0.05	0.02	0.02	0.02	0.02	0.01	0.01	-																			
44	**Ecomp475.**	0.02	0.02	0.02	0.02	0.05	0.02	0.02	0.02	0.02	0.02	0.01	0.00	-																		
45	**Ecomp479.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	-																	
46	**Ecomp484.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	-																
47	**Ecomp485.**	0.00	0.00	0.00	0.00	0.04	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	-															
48	**Ecomp495.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	-														
49	**Ecomp496.**	0.02	0.02	0.02	0.02	0.05	0.02	0.02	0.02	0.02	0.02	0.02	0.00	0.01	0.02	0.02	0.02	0.02	-													
50	**Ecomp497.**	0.02	0.02	0.02	0.02	0.05	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.00	0.02	0.02	0.02	0.02	0.00	-												
51	**Ecomp498.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.01	0.03	0.02	-											
52	**Ecomp499.**	0.00	0.00	0.00	0.00	0.04	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.01	0.03	0.02	0.01	-										
53	**Ecomp500.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	-									
54	**Ecomp501.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.01	0.03	0.02	0.01	0.00	0.00	-								
55	**Ecomp502.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.01	0.03	0.02	0.01	0.00	0.00	0.00	-							
56	**Ecomp503.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	-						
57	**Ecomp505.**	0.02	0.02	0.02	0.02	0.05	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.00	0.02	0.02	0.02	0.02	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.02	-					
58	**Ecomp506.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	-				
59	**Ecomp507.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	-			
60	**Ecomp508.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	-		
61	**Ecomp509.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	-	
62	**Ecomp510.**	0.00	0.00	0.00	0.00	0.05	0.01	0.02	0.00	0.00	0.01	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	-

Table 3.5. Population pairwise F_{ST}. Asterisk (*) indicates significant differences between sites.

	Richland1	**Sandy**	**Polecat**	**Richland2**	**Fork**	**Brush**	**Mary**	**Potomac**
Richland1	0.00	0.06	0.18 *	-0.04	0.05	0.02	0.03	0.33 *
Sandy	0.06	0.00	0.08	0.06	0.05	0.03	0.08 *	0.43 *
Polecat	0.18 *	0.08	0.00	0.30 *	0.25 *	0.18 *	0.28 *	0.65 *
Richland2	-0.04	0.06	0.30 *	0.00	0.03	0.03	0.04	0.63 *
Fork	0.05	0.05	0.25 *	0.03	0.00	0.03	0.04	0.46 *
Brush	0.02	0.03	0.18 *	0.03	0.03	0.00	-0.02	0.46 *
Mary	0.03	0.08 *	0.28*	0.04	0.03	-0.02	0.00	0.40 *
Potomac	0.33 *	0.43 *	0.65 *	0.63 *	0.46 *	0.46 *	0.40 *	0.00

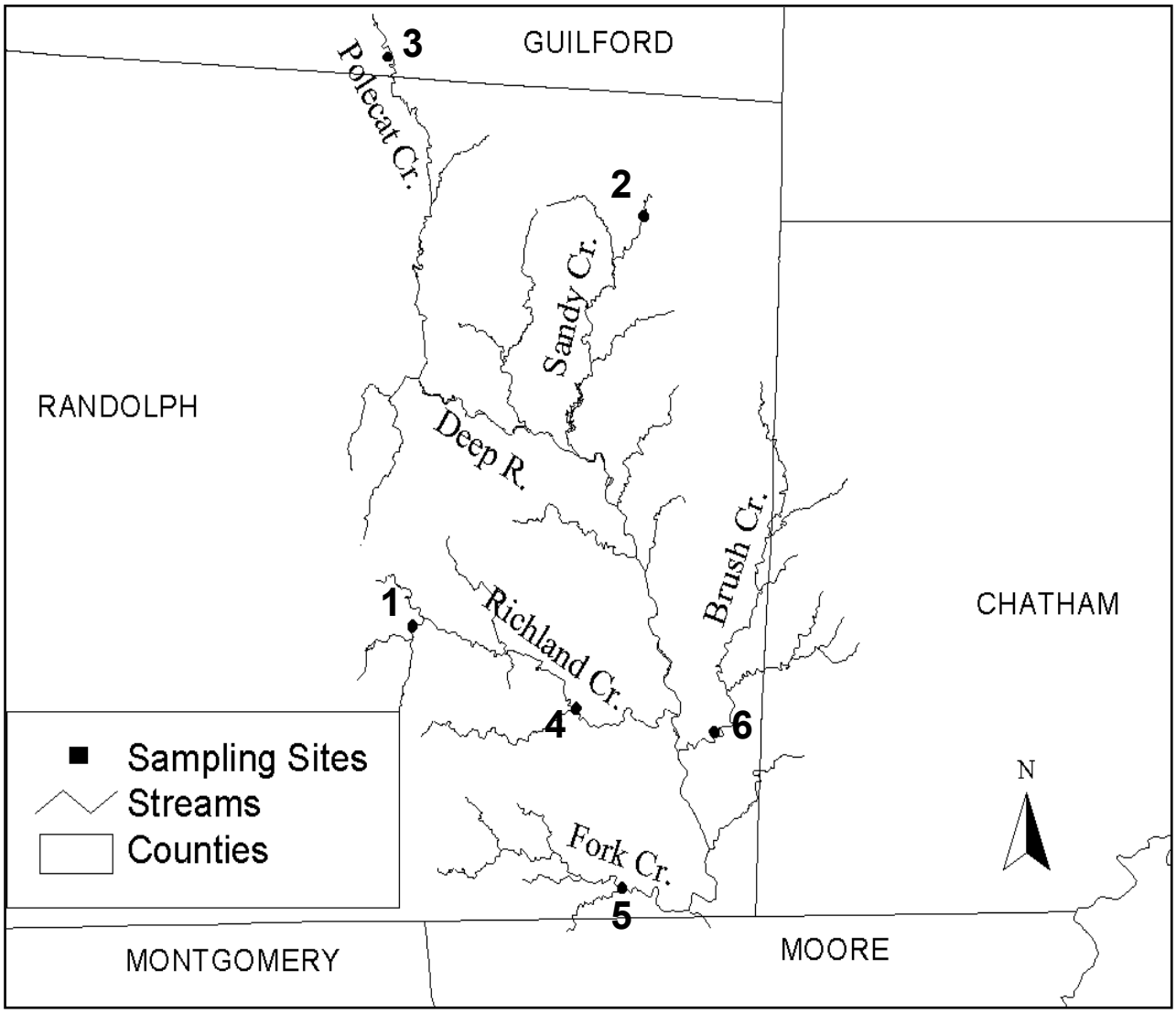


Figure 3.1. Sample site location.

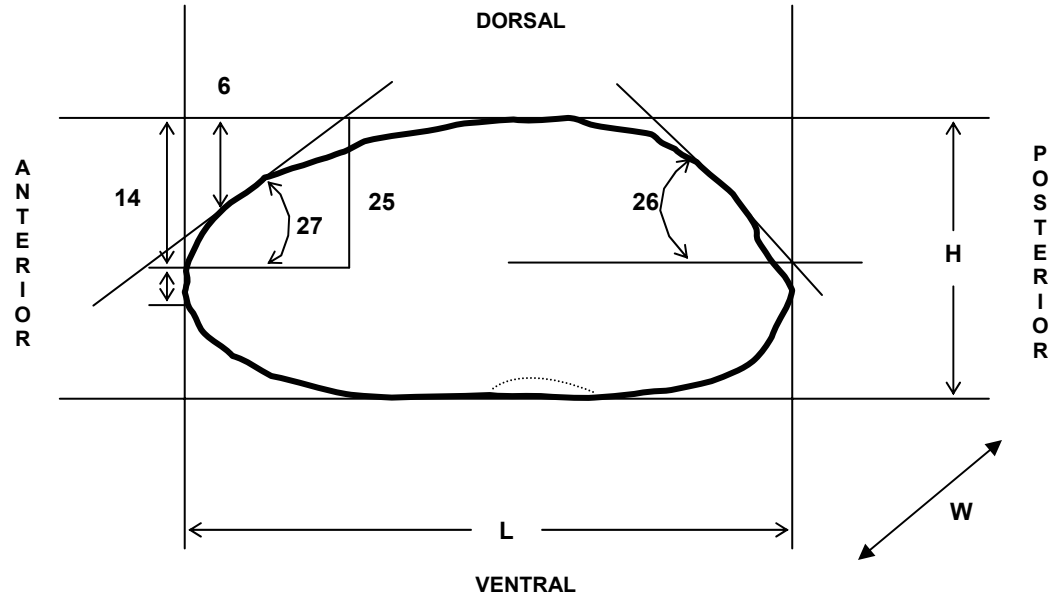


Figure 3.2. Orientation of morphometric variables included in multivariate analyses. Variable description is found in Table 3.2.

Figure 3.3. Heirarchical cluster diagram of *Elliptio complanata* collected from Cape Fear River basin, based on seven morphometric variables. Individuals in bold letter and highlighted with the letter "T" were topotype specimens.

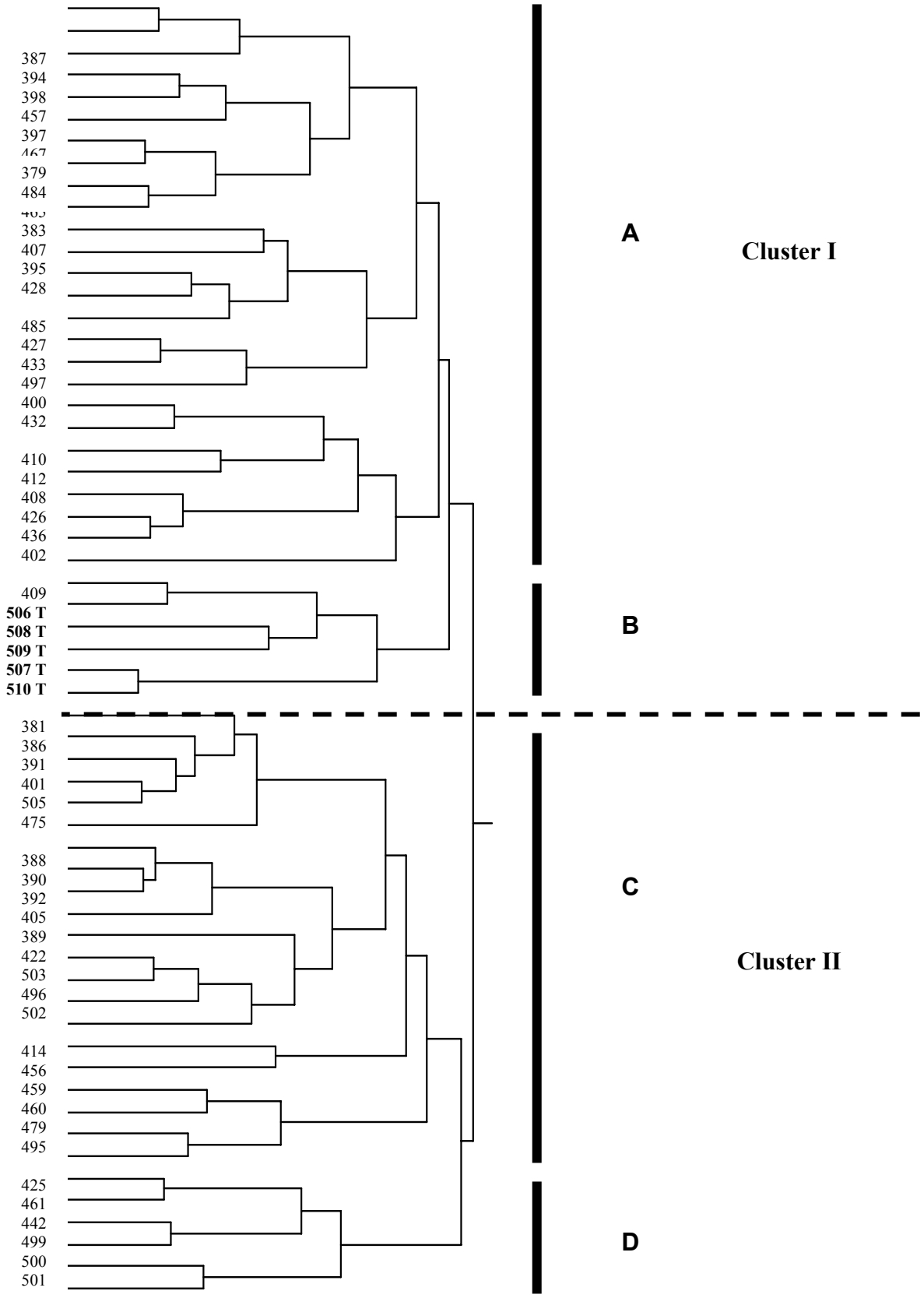
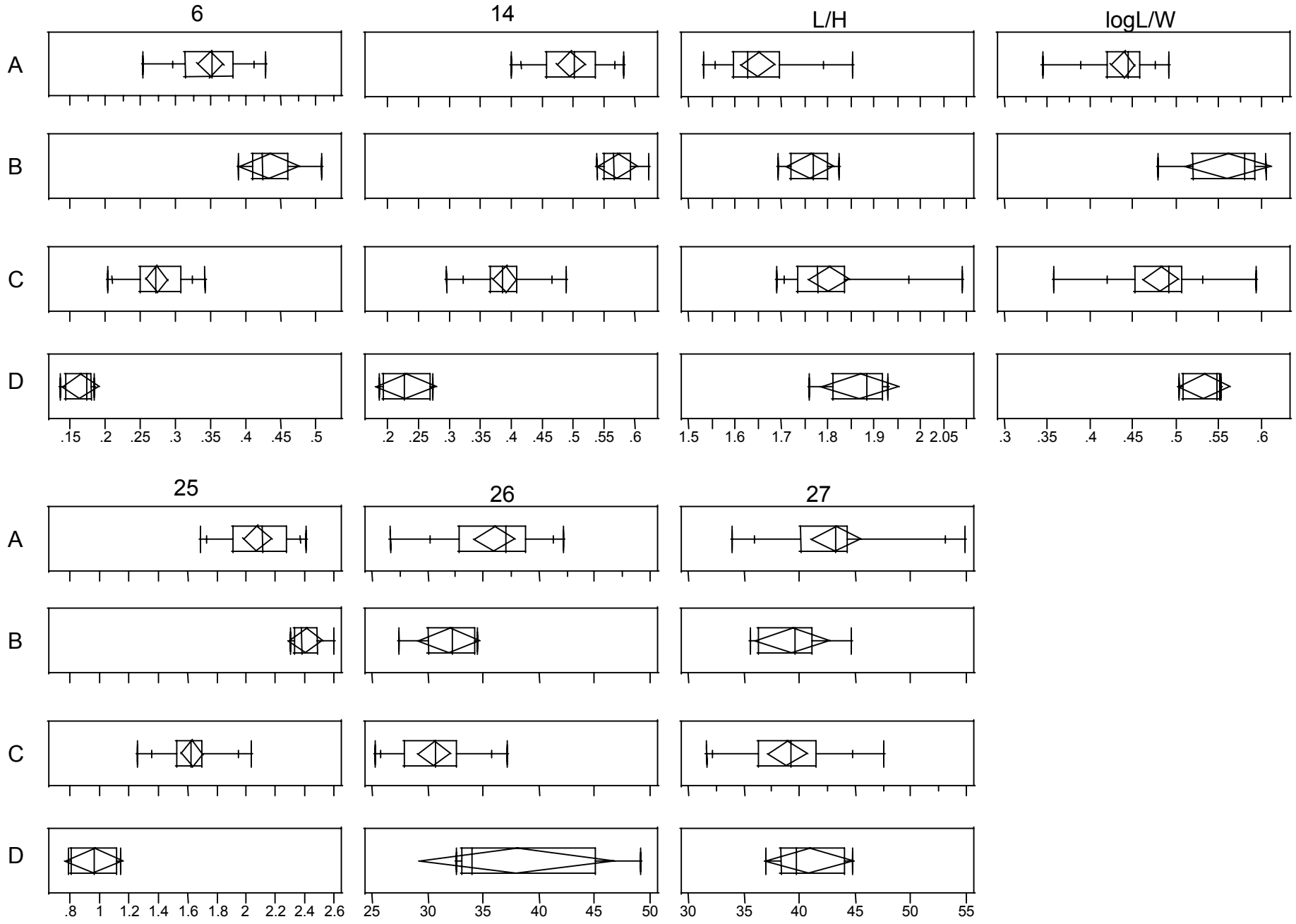

387
394
398
457
397
379
484
383
407
395
428
485
427
433
497
400
432
410
412
408
426
436
402
409
506 T
508 T
509 T
507 T
510 T
381
386
391
401
505
475
388
390
392
405
389
422
503
496
502
414
456
459
460
479
495
425
461
442
499
500
501
A
B
C
D
Cluster I
Cluster II

Figure 3.4. Summary statistics of seven morphometric variables by cluster groups. Groups were based on cluster groups from Figure 3.3. Bars from left to right represent 0, 15, 25, 50, 75, 95, and 100 % confidence interval of samples in each population. Diamonds are sample means.


6
14
L/H
logL/W
A
B
C
D
.15 .2 .25 .3 .35 .4 .45 .5
.2 .25 .3 .35 .4 .45 .5 .55 .6
1.5 1.6 1.7 1.8 1.9 2 2.05
.3 .35 .4 .45 .5 .55 .6
25
26
27
A
B
C
D
.8 1 1.2 1.4 1.6 1.8 2 2.2 2.4 2.6
25 30 35 40 45 50
30 35 40 45 50 55

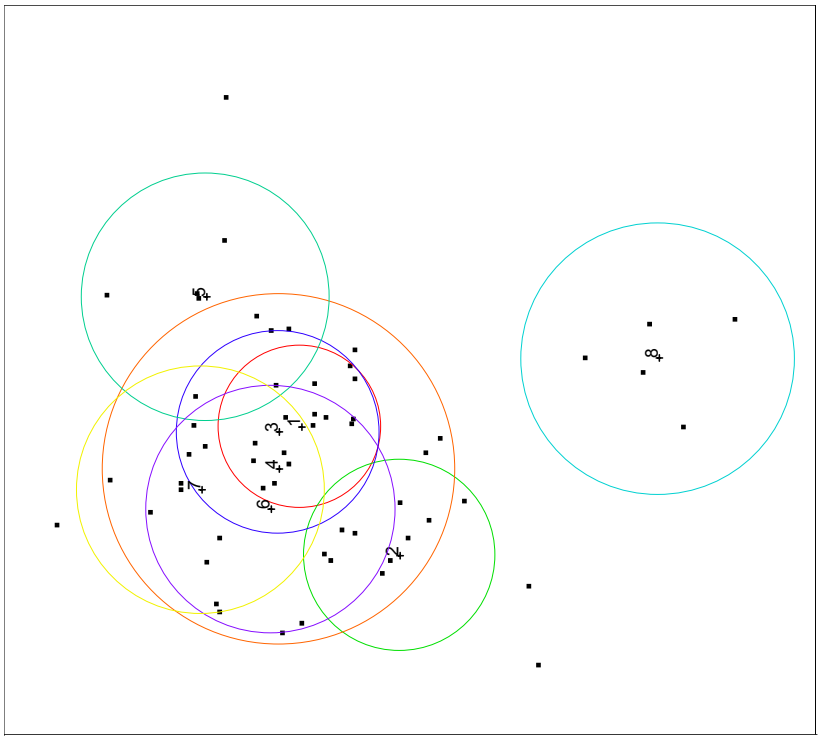


Figure 3.5. Discriminant plot of seven morphometric variables based on sites. Dots represent sample data points. Circles are 95% confidence interval for each population.

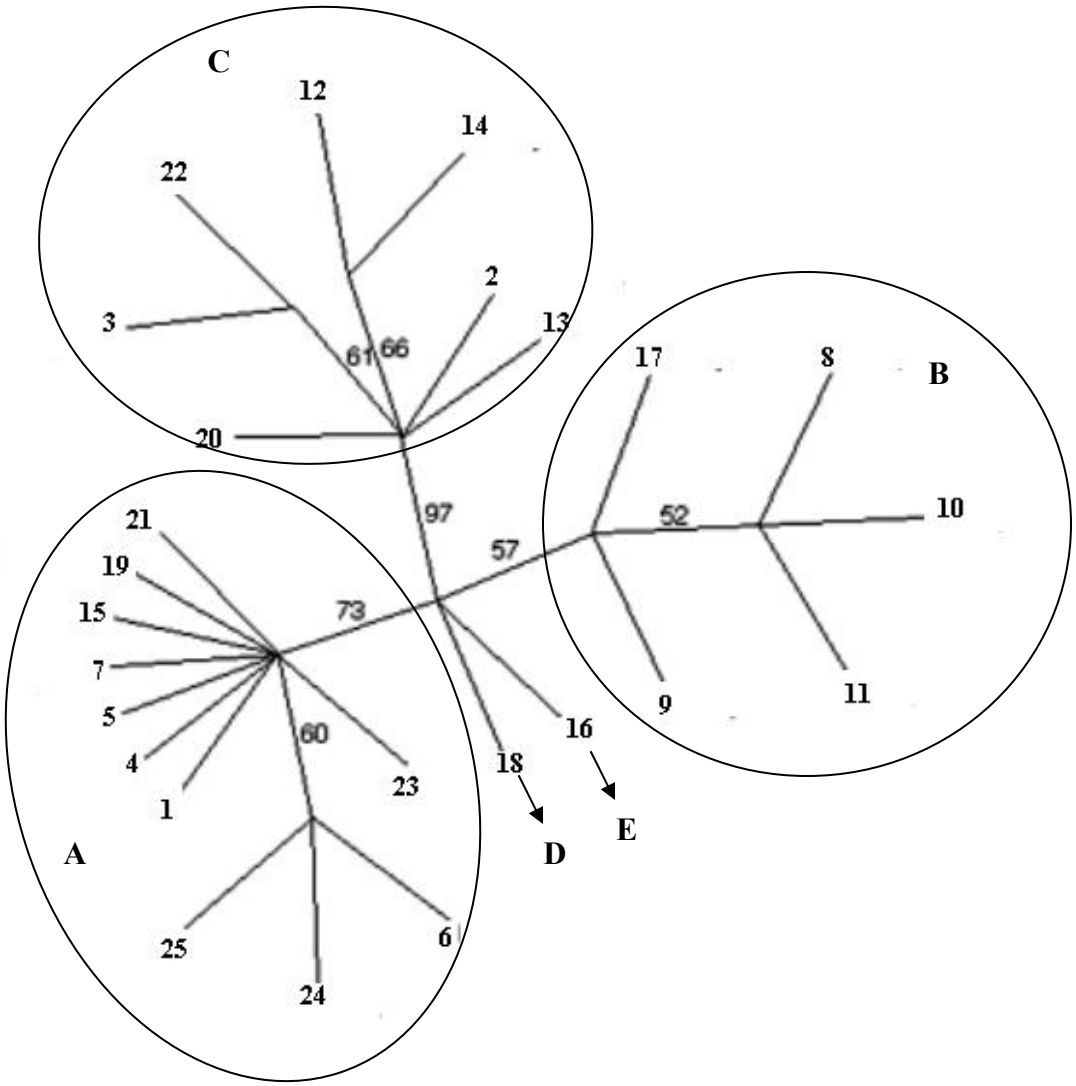


Figure 3.6. Unrooted, maximum parsimony-based dendrogram of cytochrome oxidase I (COI) sequence data for *Elliptio complanata*. Sample identification refers to haplotypes. Corresponding sample numbers are found in Table 3.2. Numbers along branches are bootstrap scores.

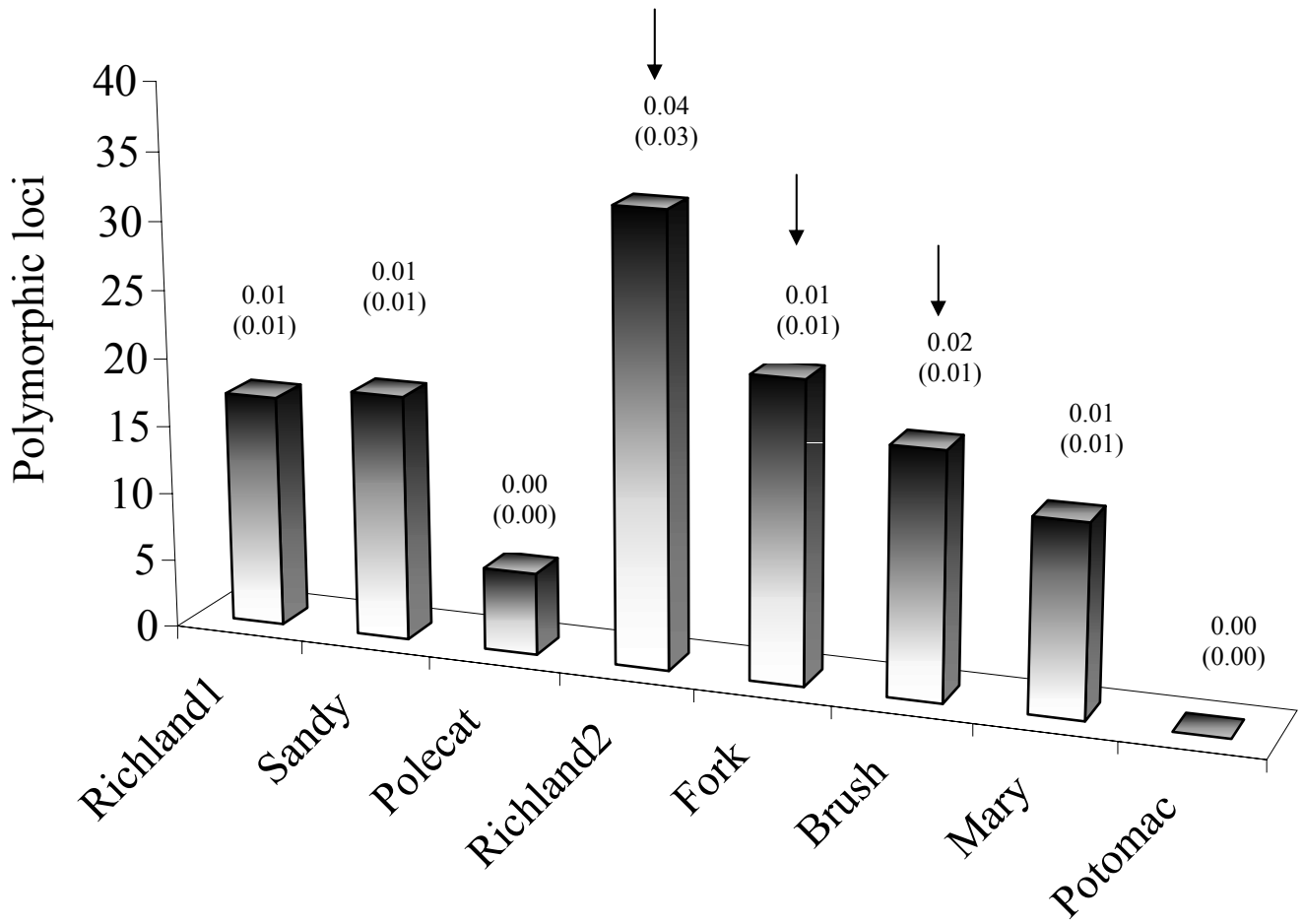


Figure 3.7. Genetic diversity indices calculated for the different sites. Bars refer to numbers of polymorphic loci. Values above bars are nucleotide diversity (s.d.). Arrows indicate downstream sites.

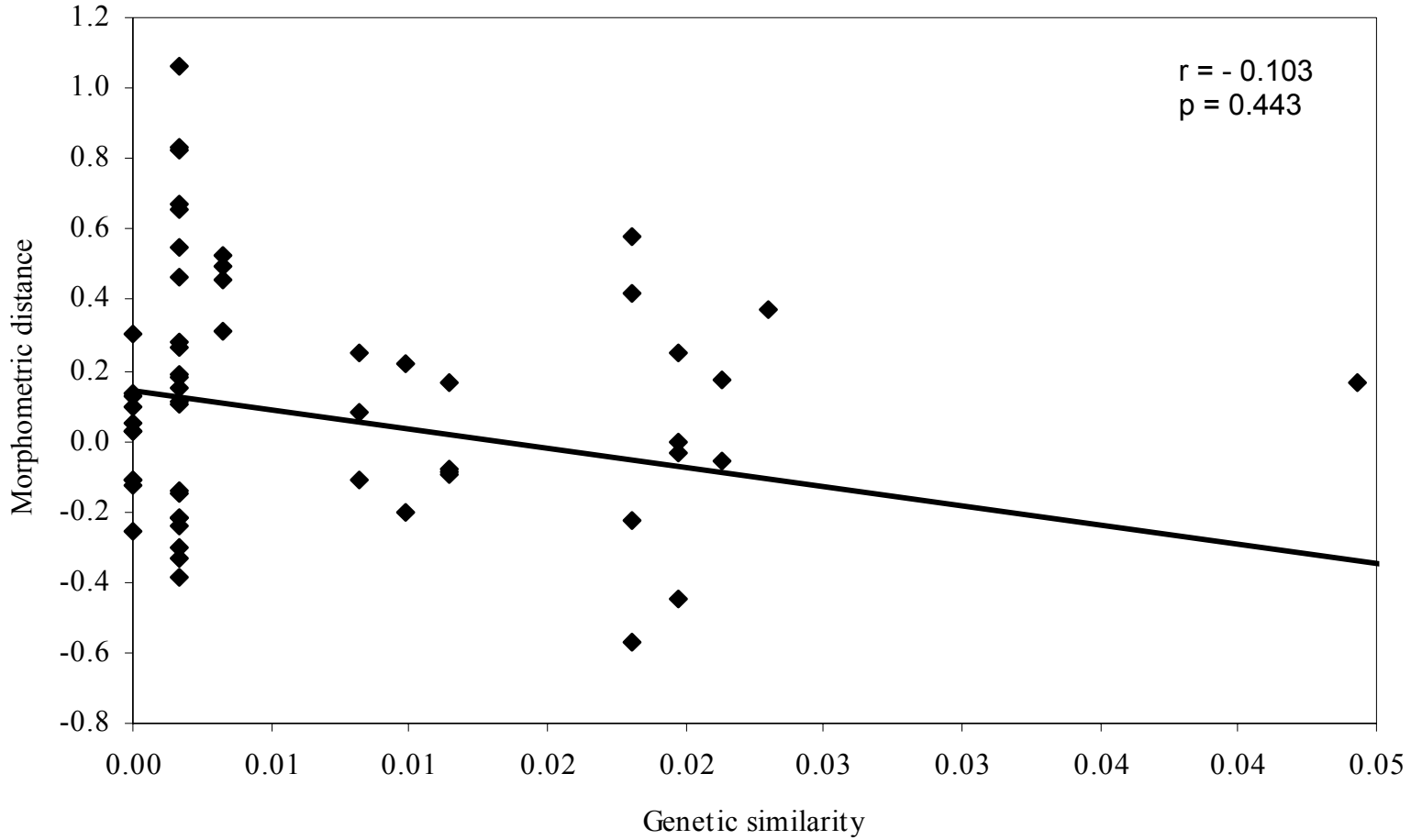


Figure 3.8. Lack of correlation between log-transformed morphometric distance and arc-sine transformed COI sequence similarity. Morphometric distance was based on Ward's distance method. Sequence data was based on pairwise similarity. Legend: r = correlation coefficient, p = p-value.

CHAPTER 4

GENETIC VARIATION WITHIN THE FRESHWATER MUSSEL, *ELLIPTIO COMPLANATA* (LIGHTFOOT, 1786): A PRELIMINARY INVESTIGATION

R. A. Molina[1], M.E. Raley[1,2], W. Gebreyes[1], B. S. Weir[3], F. Gould[4], A. E. Bogan[2], and J. F. Levine[1]

[1]Population Health and Pathobiology department, College of Veterinary Medicine, North Carolina State University, 4700 Hillsborough St., Raleigh, NC 27606, Tel. #: (919) 513-6302. Email: rmolina@unity.ncsu.edu, revmolina@yahoo.com, Jay_Levine@ncsu.edu, wagebrey@unity.ncsu.edu

[2]North Carolina Museum of Natural Sciences, Research Laboratory, 4301 Reedy Creek Road, Raleigh, NC 27607, Tel. #: (919) 733-7450 ext 753. Email: morgan.raley@ncmail.net, arthur.bogan@ncmail.net

[3] Bioinformatics Center, Statistics, 1501 Partners Building II, Box 7566, Tel. : (919) 515-3574. Email: weir@stat.ncsu.edu

[4] Department of Entomology, Box 7634, North Carolina State University, Raleigh, NC, Tel. #: (919) 515-1647. Email: fred_gould@ncsu.edu.

ABSTRACT

Populations of the common freshwater mussel, *Elliptio complanata*, from Cape Fear River (NC) were identified based on similarity of cytochrome oxidase I (COI) sequence data. Valid *E. complanata* species were then analyzed using amplified fragment length polymorphism (AFLP) to provide a genome-wide means of assessing diversity. Different estimates of genetic diversity such as heterozygosity (H), polymorphism (P), θ^B and G_{ST}-B, were calculated from the presence/absence data derived from AFLP fingerprints. In addition, fingerprint profiles were also analyzed and compared using two similarity coefficients. The

Pearson-product moment correlation coefficient and the Jaccard coefficient were used to identify the number of taxon groups existing within valid *E. complanata*. Using the Bionumerics program, cluster analysis was performed and compared using the neighbor-joining (NJ) and unweighted pair-group method using arithmetic averages (UPGMA) techniques.

Regardless of the similarity coefficient or cluster procedure, the three *E. complanata* taxon groups were always distinct. This result supports prior studies suggesting the occurrence of more than one species group in *E. complanata*. The highest bootstrap support for the three taxon groups were generated from Jaccard-based trees, indicating the procedure's utility for analyzing dominant marker-based characters. Heterozygosity values within the sampled populations within the Cape Fear River demonstrate, in part, directional gene flow structuring genetic variation, a pattern consistent with results from COI data. High H and P values compared to earlier work, suggest high gene flow among CFR populations. The results of this study strengthen earlier reports recognizing high genetic variability within *E. complanata*.

INTRODUCTION

Until recently species conservation efforts have been primarily dependent on the availability of information about a species' abundance, density, geographic distribution and life history. However, molecular genetic techniques now provide a powerful new tool to assist conservation biologists in the development of effective species preservation strategies (Berg et al. 1998, Zawko et al. 2001). Mitochondrial DNA (mtDNA) analysis has provided

field biologists with the ability to readily differentiate between species with similar morphologic and anatomic appearance. Indeed sequencing of mtDNA, cytochrome oxidase I (COI) and/or 16S rDNA have been used to distinguish between species of marine bivalves, fish, and other aquatic species (Gatt et al. 2002, Park and Foighil 2000, Staton et al. 1997).

When species abundance and geographic distribution decline, conservation biologists must heighten their concern about the genetic potential of a species to sustain itself within a specific habitat. The more diverse a species gene pool the greater its ability to adapt to environmental change (Booy et al. 2000). Clear examples are reflected in the plight of the cheetah and the black-footed ferret, which have both been brought to the brink of extinction (Amos and Balmford 2001). Although freshwater mussels receive less attention, the majority of mussel taxa are similarly imperiled (Williams et al. 1993). Limited work has been conducted to examine genetic diversity in freshwater mussel populations (Mulvey et al. 1997, Berg et al. 1998, Johnson et al. 1998).

Allozyme electrophoresis and starch gel electrophoresis have been used to study genetic variation in both marine and freshwater bivalves (Berg et al. 1998, Johnson et al. 1998, Moraga et al., 1994). However, the information generated lacked clarity and consistency (Mueller and Wolfenbarger 1999). Microsatellite analysis has been suggested as an alternative and has been used effectively for assessing genetic variability in several terrestrial species as well as playing a role in assessing paternity (Beaumont et al. 1999, Gerber et al. 2000). However initial attempts to develop nucleotide arrays for use with freshwater mussel species have been less effective (King et al. 1999). Di-nucleotide microsatellite markers with limited variability have been isolated from nuclear genes, but assessment of genetic diversity using these has not been fruitful (King et al. 1999). The

identification of useful microsatellite markers has often been time consuming and expensive (Mueller and Wolfenbarger 1999, Queller et al. 1993). Amplified fragment length polymorphism (AFLP) was a viable alternative that was more time efficient and less costly (Mueller and Wolfenbarger 1999). The technique has been used extensively to examine genetic variability in bacteria, plants, insects, and various vertebrates (Giannasi et al. 2001, Jansen et al. 1996, Marriette et al. 2001, Ravel et al. 2001). In this study, AFLP was used to assess gene pool diversity of the cosmopolitan freshwater mussel, *Elliptio complanata*. The procedure provided another tool for the study of freshwater mussel genetic diversity that facilitated sampling of sequences from the whole genome, in contrast to earlier work based on the study of genetic variation in a single gene.

METHODOLOGY

Sample acquisition

Individual samples for this study came from a subset of samples from a previously conducted cytochrome oxidase I (COI) gene sequencing analysis on *E. complanata* from the Cape Fear River, NC (CFR) and the Potomac River (Washington, D.C.) (Figure 4.1). The site was previously described in Molina and co-workers (chapters one and two). Sites one to seven were located along the CFR and the eighth site contained topotypic specimens (Potomac River), which became the basis for genetic comparison with CFR specimens (Table 4.1).

AFLP analysis

Samples used for AFLP were taken from valid species of *E. complanata* identified based on their COI similarity with topotypic speciments (chapter three). Unrooted maximum parsimony based phylogeny of COI sequenced samples yielded three groups that were well supported by bootstrap analysis (Figure 4.2). Samples belonging to the group that closely resembled topotypic specimens (group A) were then utilized for genomic fingerprinting to further investigate the purported *E. complanata* species group (Figure 4.2).

Mussel tissues of samples belonging to group A (n = 38) were collected and DNA was extracted as previously described in Molina and co-workers (chapter three). Initial genotyping attempts were conducted twice on eight individuals to assess repeatability associated with the PCR and AFLP visualization procedures. The AFLP procedure could be deemed repeatable if paired samples from each individual in the group of eight were tightly clustered based on ≥ 70% similarity (Hane et al., 1993).

The AFLP procedure on all samples was conducted following protocols adapted from an AFLP manual (LI-COR Biosciences, Lincoln, Nebraska) and Myburg and Remington (2000). Specifically, the restriction, ligation and pre-amplification steps described in the manual were followed without modification. Briefly, extracted mussel DNA was restricted with *Eco*RI and *Mse*I specific cutters then ligated with adapters to produce amplification templates. *Eco*RI and *Mse*I primers with additional A and C bases, respectively, were added to the amplicons to produce pre-selective PCR products. For the final amplification, we used three *Eco*RI/*Mse*I primer combinations (selective nucleotides AAG/CAC, AAC/CAC, ACC/CAT) to generate markers ranging in size from 50 to 700 bp. This final selective amplification step was adapted from Myburg and Remington (2000) where the polymerase

chain reaction (PCR) products from the pre-amplification step were selectively amplified using PCR in a 20 μl total volume containing: 2 μl 10X PCR buffer, 1.6 μl dNTP mixture of 2.5 mM each base (QIAGEN, Germantown, Maryland), 0.83 μl *Eco*RI primer (LI-COR Biosciences, Lincoln, Nebraska), 5 μl *Mse*I primer (LI-COR Biosciences, Lincoln, Nebraska), 0.24 μl*Thermus aquaticus* (*Taq*) polymerase of 5 U/μl (QIAGEN, Germantown, Maryland), 5.33 μl double-distilled H_20 and 5 μl of pre-amplified template DNA. The final amplification was accomplished through PCR performed on a PTC 100 thermal cycler (MJ Research Inc., Waltham, Massachusetts) using a program optimized by Myburg and Remington (2000). Then, selectively amplified PCR products were placed on a speedvac until just about dry (~1 h) to increase its concentration. Five microliters of blue stop solution (LI-COR Biosciences, Lincoln, Nebraska) were then added to terminate the remaining reaction. After denaturing at 90^{o}C for 3 min, the amplified restriction fragments were electrophoresed on denatured polyacrylamide gels and run on a LI-COR IR^2 automated sequencer (model 4200) using the e-Seq^{TM} software (version 3) (LI-COR Biosciences, Lincoln, Nebraska). The resulting picture file (in tiff format) was imported in the Bionumerics software (Applied Maths BVBA, Kortrijk, Belgium) where AFLP fingerprint profiles were analyzed using the (1) curve-based protocol of the Pearson product-moment correlation coefficient (*r*) and (2) band-based option, Jaccard coefficient. The Pearson-product moment correlation coefficient provided an indication of the relationship and strength of densitometric curves from two samples that were compared for similarity (Hane et al. 1993). The Jaccard coefficient allowed for cluster formation among genetic bands having the greatest frequency of occurrence (Jackson et al. 1989). Using these coefficients, Bionumerics made use of all polymorphic markers (+/-) from each primer pair to calculate a

matrix, which allowed for the cluster analysis of populations using neighbor-joining (NJ) and unweighted pair-group methods using arithmetic averages (UPGMA) procedures. Dendrograms from these clustering procedures were compared and particular relationships between trees were investigated to determine suitable taxon groupings resembling those found in a consensus tree (Quicke 1993). The strength of internal nodes was assessed by bootstrapping over loci with 1000 pseudoreplicates.

Analysis of genetic diversity

Percent polymorphic loci (P) were calculated from presence/absence data (+/-) as one of the measures of genetic diversity. In addition, presence/absence data were nexus formatted for running in the freeware program, Hickory v1.0 (Holsinger and Lewis 2003) to calculate additional estimates of genetic diversity. Specifically, the program allowed for calculation of heterozygosity to estimate genetic diversity, θ^B to estimate of F_{ST}, which is an estimate of population differentiation, and G_{ST}-B to determine the Bayesian estimate of Nei's G_{ST}, an equivalent of Wright's F_{ST} (Holsinger and Lewis 2003). Although the program calculated these estimates, caution is suggested in interpreting the results due to the small sample size used in this study and lack of comparison with other estimates for E. complanata using other procedures (i.e. microsatellite and allozyme), which could lead to erroneous estimates when using AFLP dominant markers (Holsinger and Lewis 2003).

RESULTS

High molecular weight DNA (size = 7987 to 9416 bp) was extracted from all the samples. Preliminary attempts at genotyping showed low noise for most samples (Figure

4.3). The first primer contained two samples (401 and 422) that were not successfully resolved and one sample (506) that had 0% similarity. The rest of the samples had similarity ranging from 75% to 100% (Figure 4.3). Band similarity for the second primer ranged from 80% to 92%, except for one unresolved sample (433) (Figure 4.3). Two samples from the third primer were resolved but showed low similarity (422 and 433). Band similarity for the rest of the samples from the third primer ranged from 80% to 98% (Figure 4.3). These results demonstrated the repeatability of the technique signifying its possible application to other freshwater mussel species.

Genetic cluster analyses of AFLP markers consistently showed variability within the topotypes (Figures 4.4 to 4.7). Regardless of procedure, three of the five topotypes were always found in the same group and the two remaining specimens exhibited high variation by grouping with other taxa. Visual inspection of the topotypes revealed differences in shell form (Figure 4.9). The three topotypes found in group I had similar forms compared to the two that belonged to group II. Although 506 and 510 were both in group II, they did not have high similarity coefficient regardless of the type of cluster procedure (Figures 4 to 7).

Despite literature stating that Pearson correlation is better than band-based techniques foridentifying fingerprint markers (Hane et al. 1993), both techniques in this study, were able to arrive at a consensus that contained three groups containing the same individuals. Group I consistently contained five individuals, three of which are topotypic specimens (507, 508 and 509) (Figures 4 to 7). Genetic similarity within this group was at 38% and 48% for Pearson correlation NJ and UPGMA procedures, respectively, and 48% and 55% for Jaccard index NJ and UPGMA, respectively. Bootstrap scores for group I had lower scores for dendrograms derived from Pearson correlation compared to Jaccard coefficient (48 for NJ,

56 for UPGMA, and 94 for NJ, 95 for UPGMA, respectively). Group II included the two remaining topotypes (506 and 510) as well as five other individuals. This group showed higher bootstrap scores for Jaccard–based trees (48 and 35 for NJ and UPGMA, respectively) than Pearson-based trees (1 and 5 for NJ and UPGMA, respectively). The third group contained seven individuals that did not group with any of the topotypes, regardless of cluster procedure. Genetic similarity for this group was variable, with a high of 45% (Figure 6) and a low of 13% (Figure 7). Group III showed no bootstrap score for all the Pearson correlation-based trees and relatively higher scores for Jaccard coefficient-based NJ and UPGMA trees (53 and 23, respectively). Other individuals not belonging in the three groups showed high variability, and their location differed in every cluster procedure employed (Figures 4 to 7).

High polymorphism characterized the AFLPs from the three primer pairs (Table 4.3). Population differentiation (θ^B) ranged from 0.07 to 0.12 for the three primer pairs (± 0.01, 0.01 and 0.02 for primers 1, 2 and 3, respectively). Genetic structuring within populations (G_{ST}-β) ranged from 0.06 (primer three) to 0.10 (primer two). Average heterozygosity was generally higher for the third primer (0.31 ± 0.02) when compared to those from the other two primers (primer one = 0.19 ± 0.02; primer two = 0.19 ± 0.01). Genetic diversity within CFR sites (average heterozygosity) revealed the lowest value for site three and the highest at site one (Figure 4.8). Average heterozygosity for the topotype site had similar values with other CFR sites (Figure 4.8).

DISCUSSION

A population genetic analysis based on AFLPs was performed on individuals that had previously been identified as members of the species of *E. complanata* based on COI sequence data. The results were repeatable enough to use for population genetic analysis. The relatively high error rate, reflected in inconsistent band presence from some samples, in the two gel runs, might be due to the quality of extracted DNA from the tissue type used. Molina and co-workers (2003 poster) reported that in comparison of adductor, mantle and foot tissues, the foot tissue was the best source of DNA for genetic sequencing regardless of tissue treatments (ethanol-preserved, frozen to –20^{o}C and no treatment). We used a combination of ethanol-preserved adductor, mantle and foot tissues in this study. Therefore, future genomic studies of this kind might have better band consistency if analysis was only based on use of foot tissues.

Genomic markers revealed three taxon groups that show fidelity of grouping in dendrograms generated by different cluster techniques. The groupings found in this AFLP analysis are consistent with groupings from a prior assessment that suggested the presence of more than one species group in *E. complanata* from the southeastern U.S. (Bogan et al., 2003). Interestingly, even the supposedly monophyletic group (topotypes) contained more than one species group when genome wide analysis was utilized. Genetic diversity at the topotypic site seemed to support the dispersal of topotypes revealed by dendrograms. High heterozygosity value for site eight indicated the existence of many heterozygotes carrying different alleles thus signifying high genetic variation (Weir 1996). Compared to COI gene diversity, genomic diversity of topotypes showed higher diversity values than COI nucleotide diversity and number of polymorphic markers (Molina et al., chapter two). Upon visual

inspection after the analysis, the two topotypic specimens thatgrouped separately from the other topotypes turned out to be different morphologically (Figure 4.9). We suggest further morphological and genetic inspection of more topotype specimens to confirm or refute this finding.

The three groups of *E. complanata* showed better bootstrap support for Jaccard-based trees compared to trees based on Pearson correlation, whether the cluster procedure was based on NJ or UPGMA. The high bootstrap support from Jaccard-based trees corroborated its utility in generating similarity coefficient for dominant marker data as it excludes negative co-occurrences, just like other similarity coefficients such as Sorensen-Dice and Ochai (Jackson et al. 1998, Duarte et al. 1999). Jaccard similarity had been widely employed in cluster analysis of another dominant marker, RAPD because of ease in interpretation resulting from its simple algorithm (Duarte et al. 1999). Basically, it starts cluster formation from the frequently occurring species (or marker/loci) and allows for the exclusion of negative co-occurrences (band absence in two species as indicator of high similarity) in its similarity algorithm (Duarte et al. 1999). Jackson and co-workers (1998) proposed that similarity coefficients, such as Jaccard's, reflect a general size/shape effect for several fish species in Ontario. Unfortunately, this trend was not apparent in the present data, because groups did not display distinct shape differences. Collecting additional samples from additional sites could provide a more accurate reflection of the relationship between the size andshape of *E. complanata* pattern that they reported.

The use of genome-wide marker data corroborated in part the diversity trends generated from COI sequencing of *E. complanata*. Heterozygosity within CFR sites suggested that the most upstream site (site three) was the less diverse genetically compared to

the most downstream site (site five). This result, apparent among creeks, is consistent with *E. complanata* genetic variation based on gene sequencing results, which depicted directional gene flow structuring genetic variation (Molina et al., chapter two). Downstream sites were found to have higher genetic diversity than upstream sites for *E. complanata* in the CFR, based on the COI gene (Molina et al., chapter two). Whitehead and co-workers (2003) reported a similar finding in California populations of the freshwater fish, *Catostomus occidentalis*. Higher genetic diversity was found for *C. occidentalis* found in the downstream site compared to those in the upstream location. However, heterozygosity from the two CFR sites located in a single creek (sites one and four) did not follow the directional gene flow pattern noted previously. The more downstream site (site four) within Richland creek had lower heterozygosity than the site above it (site one). The discrepancy in genetic trend might be a factor of the small number of sites where samples were taken. Thus, we suggest that additional studies that expand the number of creeks sampled and the number of samples collected in each creek be investigated to corroborate initial findings and the COI sequencing results that suggest directional gene flow structuring genetic variation (Molina et al., chapter two).

Genomic markers corroborated and strengthened earlier claims recognizing high variability within *E. complanata*. As expected, heterozygosity (H) values based on AFLP markers yielded values that were 35% higher than H values generated from allozyme data because of the recognized capability of standard electrophoretic methods to underestimate genetic variability (Davis et al. 1981). Paired with the high polymorphism reported in this study, high H values suggest high gene flow among *E. complanata* populations from the CFR, possibly brought about by an abundance of host-fishes with wide dispersal abilities.

Management or conservation schemes for species depicting such genetic variability trend may suggest that preservation of populations from the different stream locations will conserve most of the taxon group's genetic diversity.

REFERENCES

Amos, W. and A. Balmford. 2001. When does conservation genetics matter? Heredity 87: 257-265.

Berg, D.J., E.G. Cantonwine, W.R. Hoeh and S.I. Guttman. 1998. Genetic structure of Quadrula quadrula (Bivalvia: Unionidae): Little variation across large distances. Journal of Shellfish Research 17(5): 1365-1373.

Beaumont, M.A. and M.W. Bruford. 1999. Chapter 13. Microsatellites in conservation genetics. In Microsatellites: Evolution and Applications, Goldstein, D.B. and C. Schlotterer (eds.). Oxford University Press: Oxford, England.

Bogan, A.E., J.M. Serb and C. Lydeard. 2003. Determination of the systematic position and relationships of the Elliptio complex (Mollusca: Bivalvia: Unionidae) with distributions in North and South Carolina. A Final Report submitted to USGS (Contract 01HQAG0185).

Booy, G., R.J.J. Hendriks, M.J.M. Smulders, J.M. Van Groenendael and B. Vosman. 2000. Genetic diversity and survival of populations. Plant Biology 2: 379-395.

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

Duarte, J.M., J.B. dos Santos and L.C. Melo. 1999. Comparison of similarity coefficients based on RAPD markers in the common bean. Genetics and Molecular Biology 22(3): 427-432.

Gatt, M.H., D.J. Fraser, A.P. Liskauskas and M.M. Ferguson. 2002. Mitochondrial DNA variation and stock structure of walleyes from eastern Lake Huron: An analysis of contemporary and historical samples. Transactions of the American Fisheries Society 131(1): 99-108.

Gerber, S., S. Mariette, R. Streiff, C. Bodenes and A. Kremer. 2000. Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. Molecular Ecology 9(8): 1037-1038.

Giannasi, N., R.S. Thorpe and A. Malhotra. 2001. The use of amplified fragment length polymorphism in determining species trees at fine taxonomic levels: analysis of a medically important snake, Trimeresurus albolabris. Molecular Ecology 10: 419-426.

Hane, B.G., K. Jager, H.G. Drexler. 1993. The Pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. Electrophoresis 14: 967-972.

Holsinger, K.E. and P. O. Lewis. 2003. Hickory: A Package for Analysis of Population Genetic Data v1.0.

Jackson, D.A., K.M. Somers, H.H. Harvey. 1989. Similarity coefficients: measures of co-occurrence and association or simply measures of occurrence? The American Naturalist 133(3): 436-453.

Janssen, P., R. Coopman, G. Huys, et al. 1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. Microbiology 142: 1881-1883.

Johnson, R.L., F.Q. Liang, C.D. Milam and J.L. Farris. 1998. Genetic diversity and cellulolytic activity among several species of unionid bivalves in Arkansas. Journal of Shellfish Research 17(5): 1375-1382.

King, T.L., M.E. Eackles, B. Gjetvaj and W. R. Hoeh. 1999. Intraspecific phylogeography of Lasmigona subviridis (Bivalvia: Unionidae): conservation implications of range discontinuity. Molecular Ecology 8: S65-S78.

Marriette, S., D. Chagne, C. Lezier, P. Pastuszka, A. Raffin, C. Plomion and A. Kremer. 2001. Genetic diversity within and among Pinus pinaster populations: comparison between AFLP and microsatellite markers. Heredity 86: 469-479.

Mock, K.E., T.C. Theimer, O.E. Rhodes Jr., D. L. Greenberg and P. Keim. 2002. Genetic variation across the historical range of the wild turkey (Meleagris gallopavo). Molecular Ecology 11: 643-657.

Mock, K.E., J.C. Brim-Box, M.P. Miller, M.E. Downing and W.R. Hoeh. 2004. Genetic diversity and divergence among freshwater mussel (Anodonta) populations in the Bonneville Basin of Utah. Molecular Ecology.

Moraga, D., D. Jollivet and F. Denis. 1994. Genetic differentiation across the western Pacific populations of the hydrothermal vent bivalve Bathymodiolus spp. and the Eastern Pacific (13^{o}N) population of Bathymodiolus-Thermophilus. Deep-Sea Research Part I-Oceanogr. Res. Pap. 41(10): 1551-1567.

Mueller, U.G. and L.L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. TREE 14(10): 389-394.

Mulvey, M., C. Lydeard, D.L. Pyer, K.M. Hicks, J. Brim-Box, J.D. Williams and R.S. Butler. 1997. Conservation genetics of North American freshwater mussels Amblema and Megalonaias. Conservation Biology 11: 868-878.

Myburg, A.A. and D.L. Remington. 2000. Protocol for high-throughput AFLP analysis using LI-COR IR^2 Automated Sequencers.

Park, J.K. and D.O. Foighil. 2000. Genetic diversity of Oceanic Island Lasaea (Mollusca: Bivalvia) lineages exceeds that of continental populations in the northwestern Atlantic. Biological Bulletin 198(3): 396-403.

Parmalee, P.W. & A.E. Bogan. 1998. The Freshwater Mussels of Tennessee. The University of Tennessee Press: Knoxville, 328 pp.

Queller, D.C., J.E. Strassmann and C.R. Hughes. 1993. Microsatellites and kinship. TREE 8(8): 285-288.

Quicke, D.L.J., 1993. Principles and Techniques of Contemporary Taxonomy. Chapman & Hall: Blackie Academic & Professional, Glasgow, NZ. 311 pp.

Ravel, S., N. Monteny, D. Velasco Olmos, J. Escalante Verdugo and G. Cuny. 2001. A preliminary study of the population genetics of Aedes aegypti (Diptera: Culicidae) from Mexico using microsatellite and AFLP markers. Acta Tropica 78(3): 241-250.

Staton, J.L., L.L. Daehler and W.M. Brown. 1997. Mitochondrial gene arrangement of the horseshoe crab Limulus polyphemus L: Conservation of major features among arthropod classes. Molecular Biology and Evolution 14(8): 867-874.

Weir, B.S. 1996. Genetic Data Analysis II. Methods for Discrete Population Genetic Data. Sinauer Associates: Massachusetts. 445 pp.

Willams, 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.

Zawko, G., S.L. Krauss, K.W. Dixon and K. Sivasithamparam. 2001. Conservation genetics of the rare and endangered Leucopogon obtectus (Ericaceae). Molecular Ecology 10: 2389-2396.

Table 4.1. Sample population location, description and sample sizes.

Site Number *(sample number)*	**Name**	**Location**	**Latitude**	**Longitude**	**Specimen number**
1 *(n = 10)*	Richland1	Richland creek (upstream)	35.65	-79.78	379 to 398
2 *(n = 3)*	Sandy	Sandy creek	35.79	-79.67	399 to 418
3 *(n = 8)*	Polecat	Polecat creek	35.92	-79.8	419 to 438
4 *(n = 1)*	Richland2	Richland creek (downstream)	35.61	-79.68	439 to 458
5 *(n = 1)*	Fork	Fork creek	35.53	-79.64	459 to 478
6 *(n = 5)*	Brush	Brush creek	35.6	-79.58	479 to 498
7 *(n = 5)*	Mary	Mary's creek			499 to 505
8 *(n = 5)*	Topotype	Potomac River, Washington, D.C	39.00	77.25	506 to 510

Table 4.2. List of COI haplotypes and corresponding sample numbers. COI haplotypes belonging to group A are in bold.

Haplotype	Sample number
1	**Ecomp379, Ecomp389, Ecomp400, Ecomp401, Ecomp402, Ecomp422, Ecomp425, Ecomp426, Ecomp428, Ecomp435, Ecomp436, Ecomp485**
2	Ecomp381, Ecomp391, Ecomp474, Ecomp497, Ecomp505
3	Ecomp383
4	**Ecomp386, Ecomp387, Ecomp392, Ecomp398, Ecomp442**
5	**Ecomp388, Ecomp397, Ecomp503, Ecomp506, Ecomp507, Ecomp508, Ecomp509, Ecomp510**
6	**Ecomp390, Ecomp479, Ecomp484, Ecomp500**
7	**Ecomp394**
8	Ecomp395, Ecomp465
9	Ecomp399, Ecomp405, Ecomp427, Ecomp467
10	Ecomp407, Ecomp409
11	Ecomp408
12	Ecomp410
13	Ecomp412
14	Ecomp414
15	**Ecomp432, Ecomp433**
16	Ecomp456
17	Ecomp457
18	Ecomp459
19	**Ecomp460**
20	Ecomp475
21	**Ecomp495**
22	Ecomp496
23	**Ecomp498**
24	**Ecomp499**
25	**Ecomp501, Ecomp502**

Table 4.3. Measures of genetic diversity. Refer to text for description of each parameter.

	Primer 1		Primer 2		Primer 3	
Parameters	**Mean**	**s.d.**	**Mean**	**s.d.**	**Mean**	**s.d.**
Theta-β	0.08	0.01	0.12	0.02	0.07	0.01
G_{ST}-β	0.07	0.01	0.10	0.02	0.06	0.01
% Polymorphism	0.84		0.98		0.97	

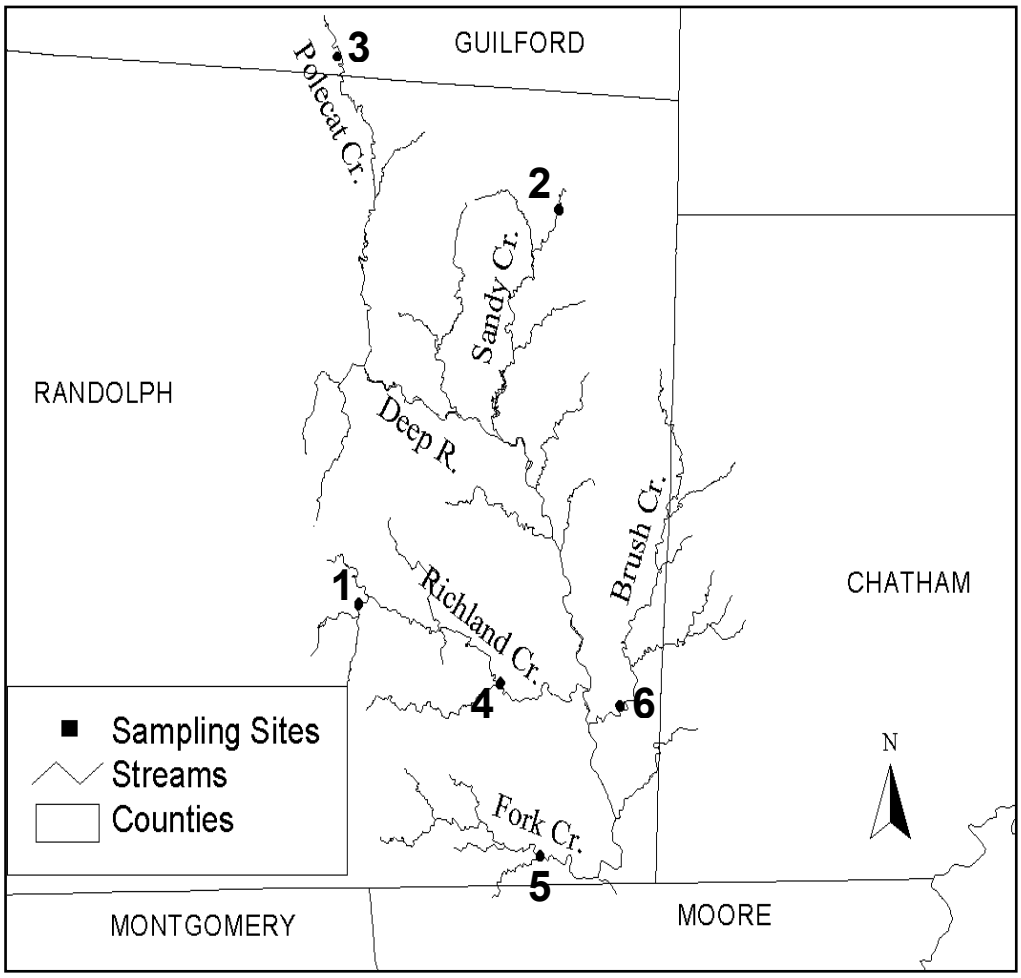


Figure 4.1. Site location.

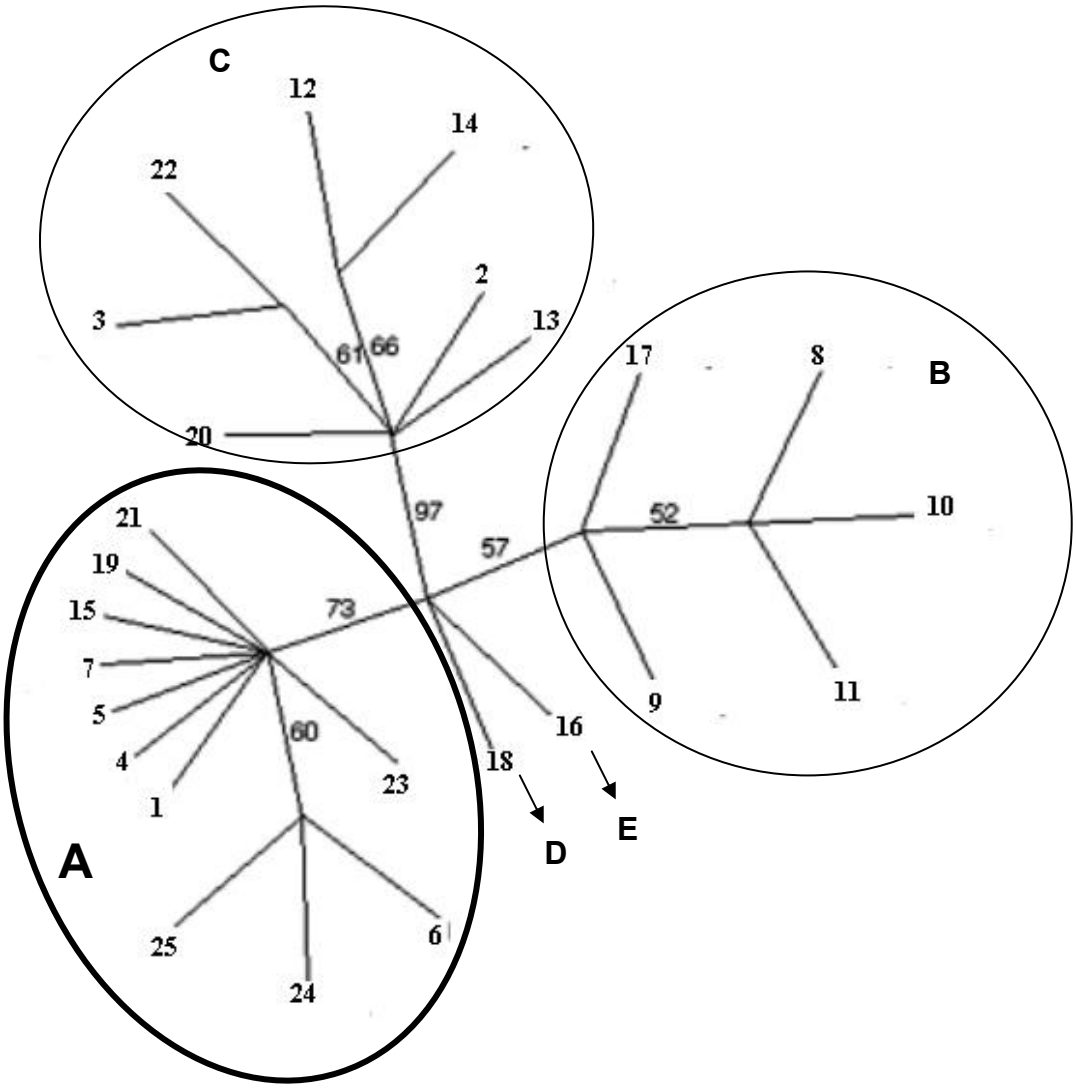
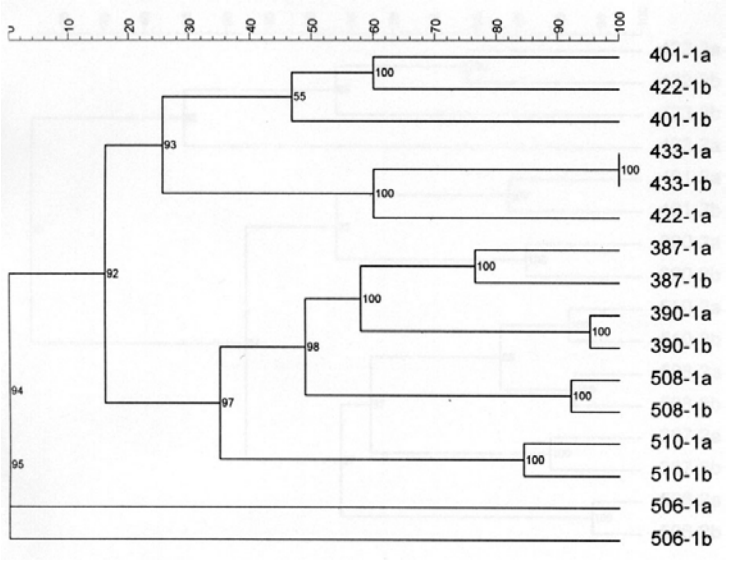
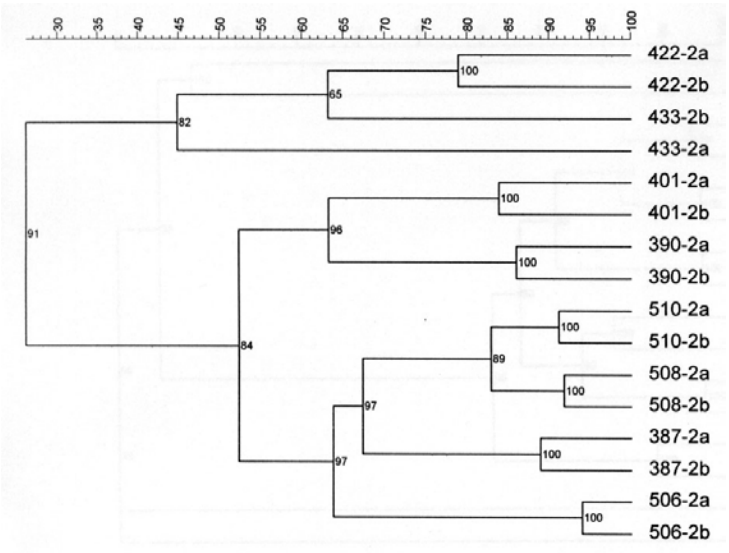


Figure 4.2. Unrooted, maximum parsimony-based dendrogram of cytochrome oxidase I (COI) sequence data for *Elliptio complanata*. Group A maybe considered valid *E. complanata* specimens due to their grouping with topotypic samples. Sample identification refers to haplotypes. Corresponding sample numbers are found in Table 4.2. Numbers along branches are bootstrap scores.

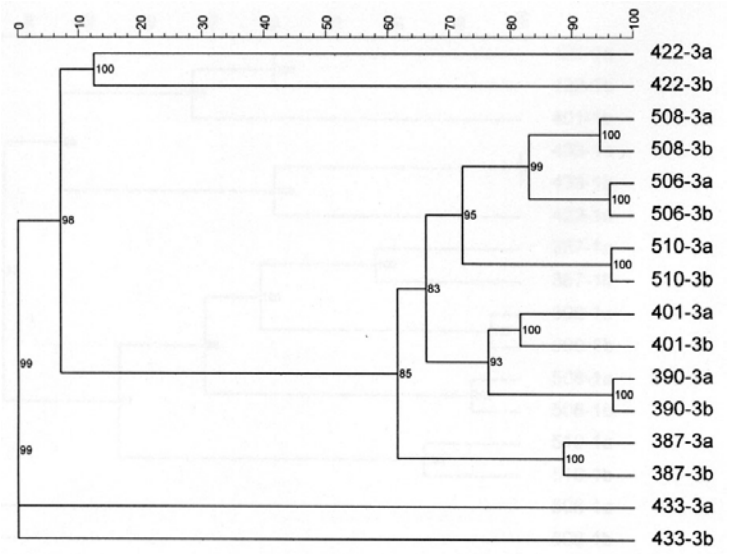
Figure 4.3. Dendrograms of the three AFLP primers (1 = first primer: AAG/CAC, 2 = second primer: AAC/CAC, 3 = third primer: ACC/CAT) taken from eight samples of valid *E. complanata* population, which were replicated (a = first replicate, b = second replicate). Values in internal nodes are bootstrap scores.


0 10 20 30 40 50 60 70 80 90 100
401-1a
422-1b
401-1b
433-1a
433-1b
422-1a
387-1a
387-1b
390-1a
390-1b
508-1a
508-1b
510-1a
510-1b
506-1a
506-1b
100
55
93
100
100
92
100
100
100
98
94
100
97
95
100

1


30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
422-2a
422-2b
433-2b
433-2a
401-2a
401-2b
390-2a
390-2b
510-2a
510-2b
508-2a
508-2b
387-2a
387-2b
506-2a
506-2b
100
65
82
100
98
91
100
100
89
84
97
100
97
100

2


0 10 20 30 40 50 60 70 80 90 100
422-3a
422-3b
508-3a
508-3b
506-3a
506-3b
510-3a
510-3b
401-3a
401-3b
390-3a
390-3b
387-3a
387-3b
433-3a
433-3b
100
100
99
100
95
98
100
83
100
93
99
85
100
99
100

3

Figure 4.4. Dendrogram of *E. complanata* based on Pearson correlation and neighbor-joining procedures. Values between internal nodes are bootstrap values (1000 permutations). Letter T refers to topotypes.

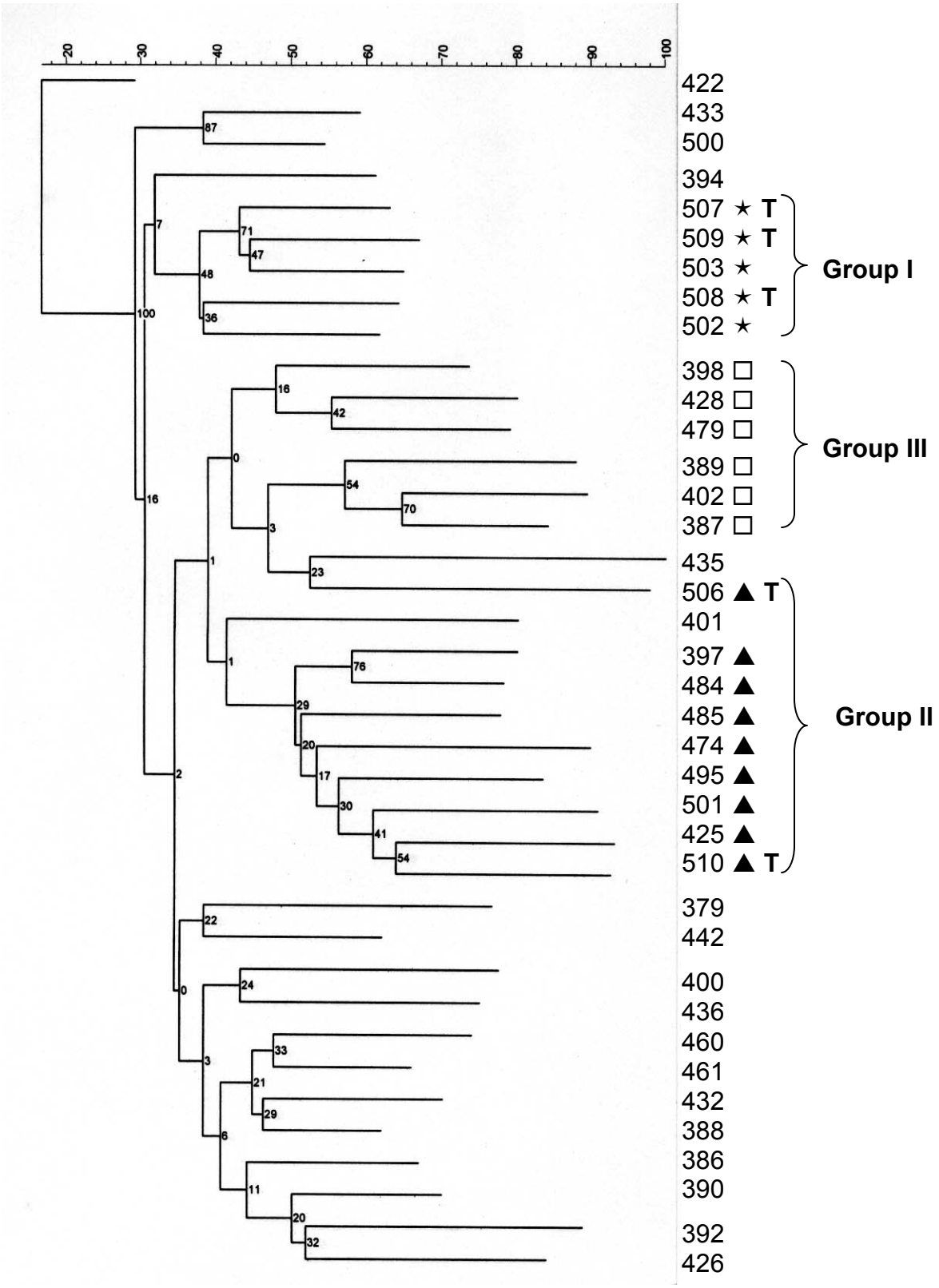

20
30
40
50
60
70
80
90
100
422
433
500
394
507 ★ T
509 ★ T
503 ★
508 ★ T
502 ★
Group I
398 □
428 □
479 □
389 □
402 □
387 □
Group III
435
506 ▲ T
401
397 ▲
484 ▲
485 ▲
474 ▲
495 ▲
501 ▲
425 ▲
510 ▲ T
Group II
379
442
400
436
460
461
432
388
386
390
392
426

Figure 4.5. Dendrogram of *E. complanata* based on Pearson correlation and UPGMA procedures. Values between internal nodes are bootstrap values (1000 permutations). Letter T refers to topotypes.

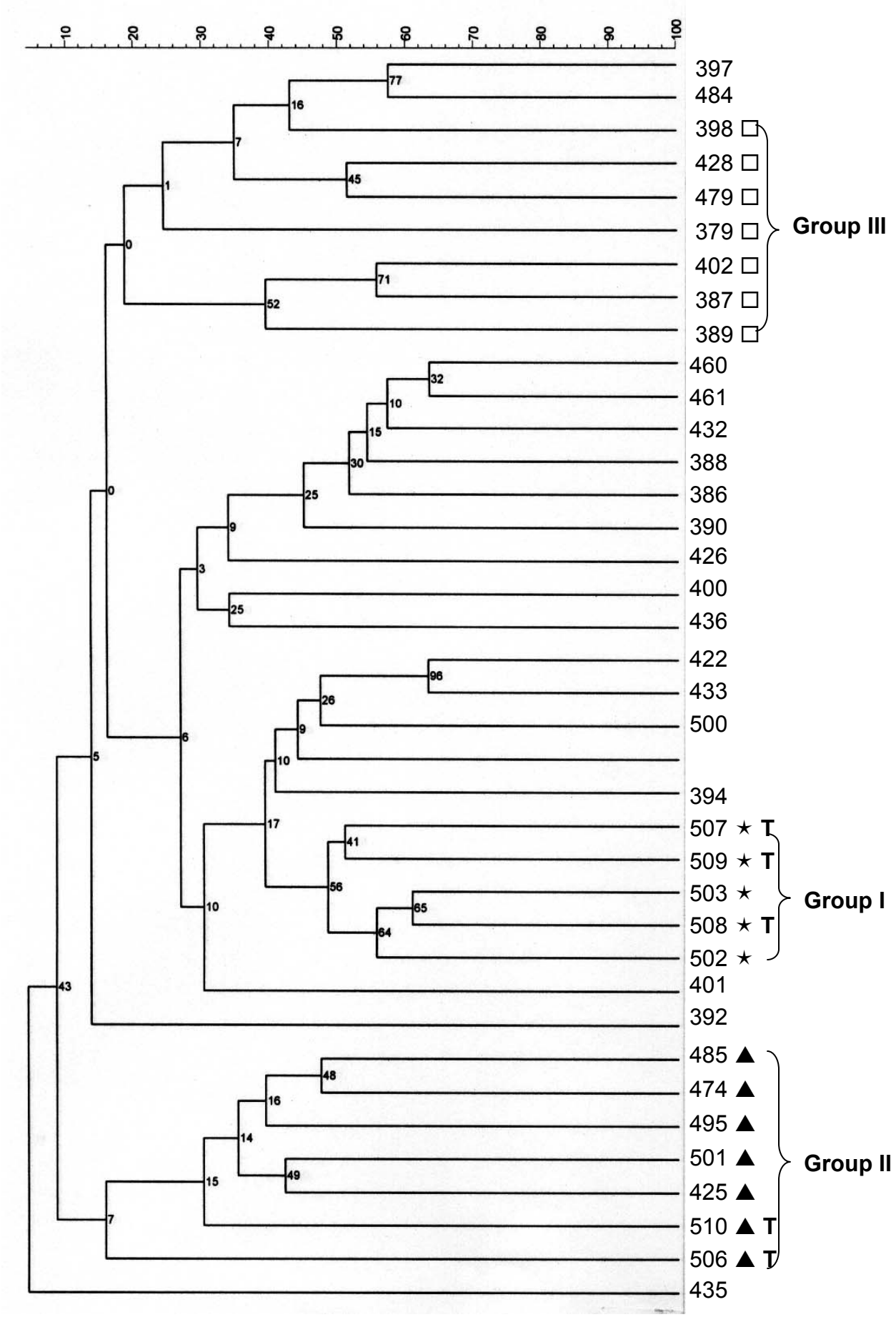

10
20
30
40
50
60
70
80
90
100
397
484
398 □
428 □
479 □
379 □
402 □
387 □
389 □
Group III
460
461
432
388
386
390
426
400
436
422
433
500
394
507 ★ T
509 ★ T
503 ★
508 ★ T
502 ★
Group I
401
392
485 ▲
474 ▲
495 ▲
501 ▲
425 ▲
510 ▲ T
506 ▲ T
Group II
435

Figure 4.6. Dendrogram of *E. complanata* based on Jaccard similarity and neighbor-joining procedures. Values between internal nodes are bootstrap values (1000 permutations). Letter T refers to topotypes.

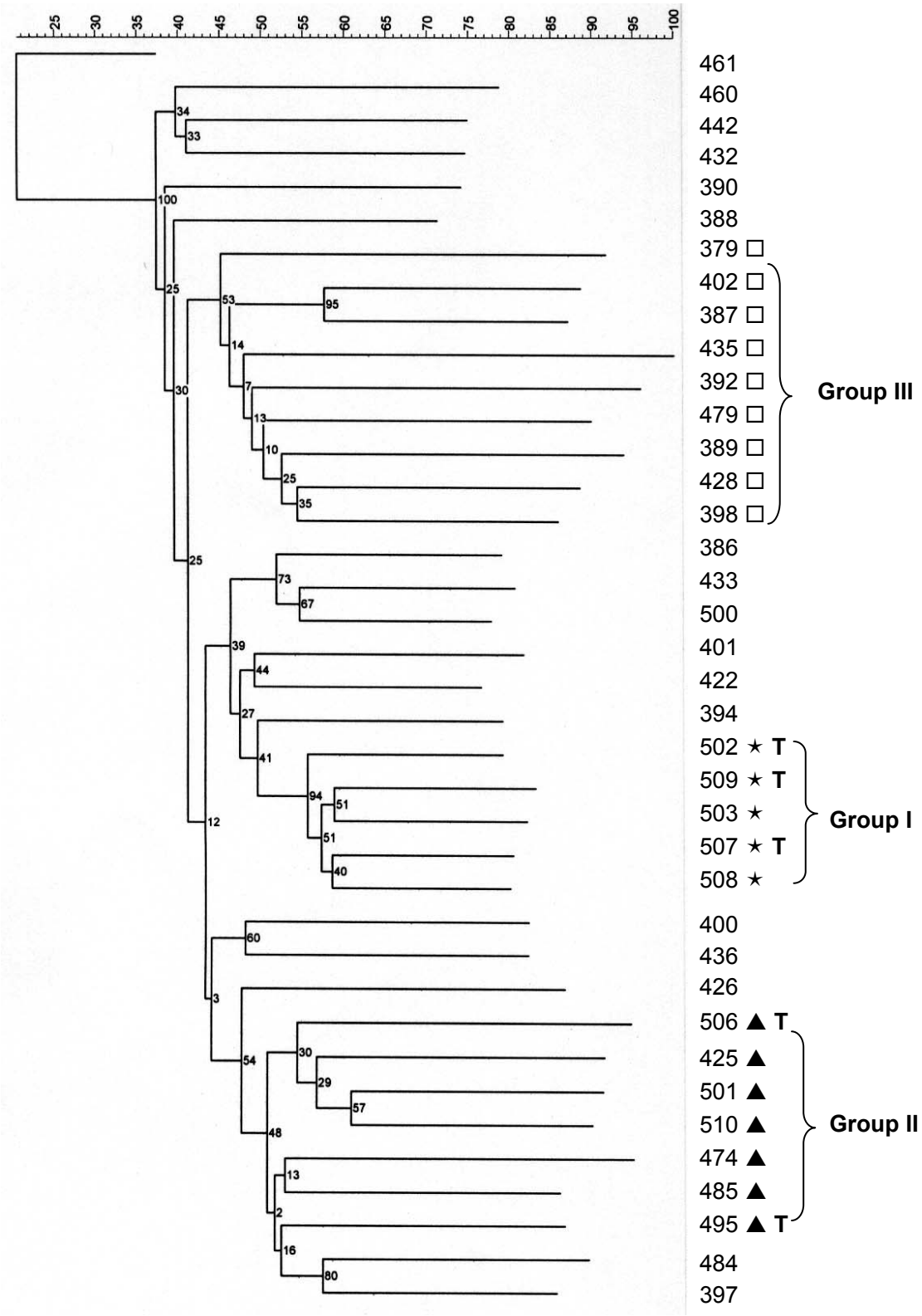
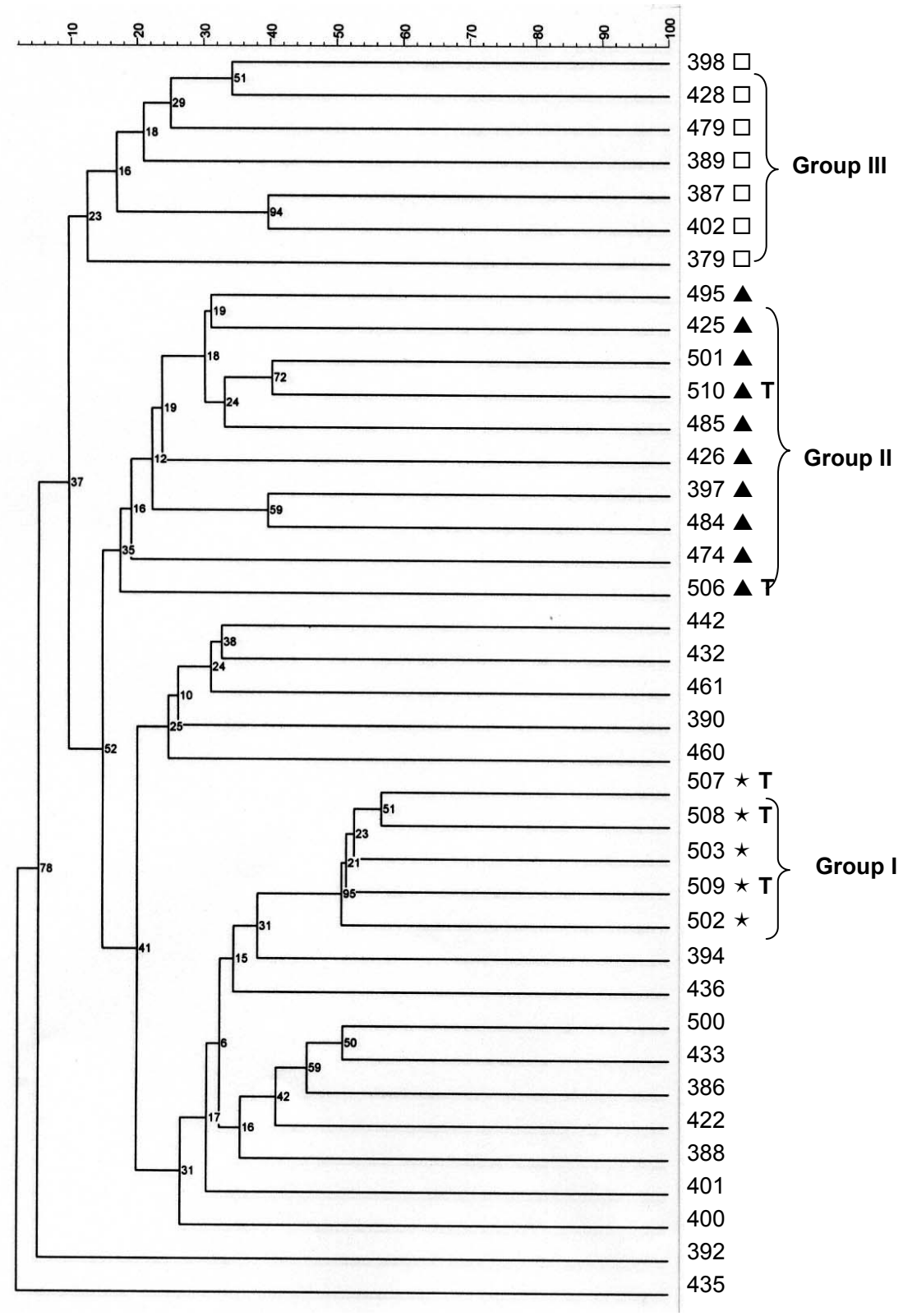

25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100
461
460
442
432
390
388
379 □
402 □
387 □
435 □
392 □
479 □
389 □
428 □
398 □
Group III
386
433
500
401
422
394
502 ★ T
509 ★ T
503 ★
507 ★ T
508 ★
Group I
400
436
426
506 ▲ T
425 ▲
501 ▲
510 ▲
474 ▲
485 ▲
495 ▲ T
Group II
484
397

Figure 4.7. Dendrogram of *E. complanata* based on Jaccard similarity and UPGMA procedures. Values between internal nodes are bootstrap values (1000 permutations). Letter T refers to topotypes.


10
20
30
40
50
60
70
80
90
100
398 □
428 □
479 □
389 □
387 □
402 □
379 □
Group III
495 ▲
425 ▲
501 ▲
510 ▲ T
485 ▲
426 ▲
397 ▲
484 ▲
474 ▲
506 ▲ T
Group II
442
432
461
390
460
507 ★ T
508 ★ T
503 ★
509 ★ T
502 ★
Group I
394
436
500
433
386
422
388
401
400
392
435

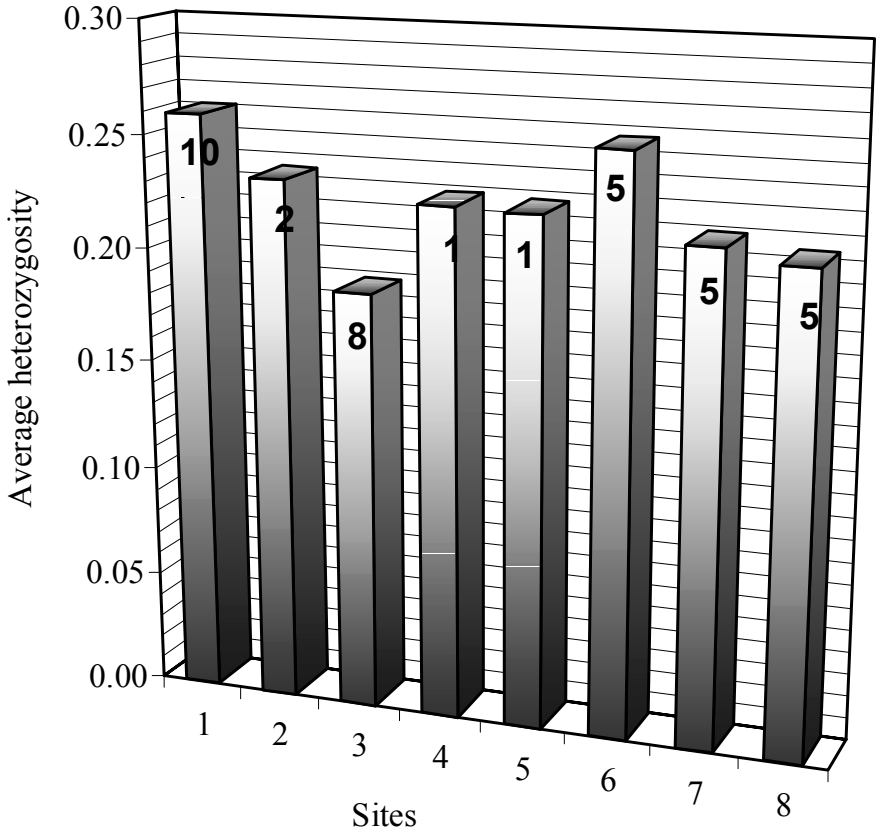


Figure 4.8. Average heterozygosity of the different sites, based on three AFLP primers. Numbers in bars are sample sizes. Sites one to seven are from CFR while site eight is from Potomac River (topotype).

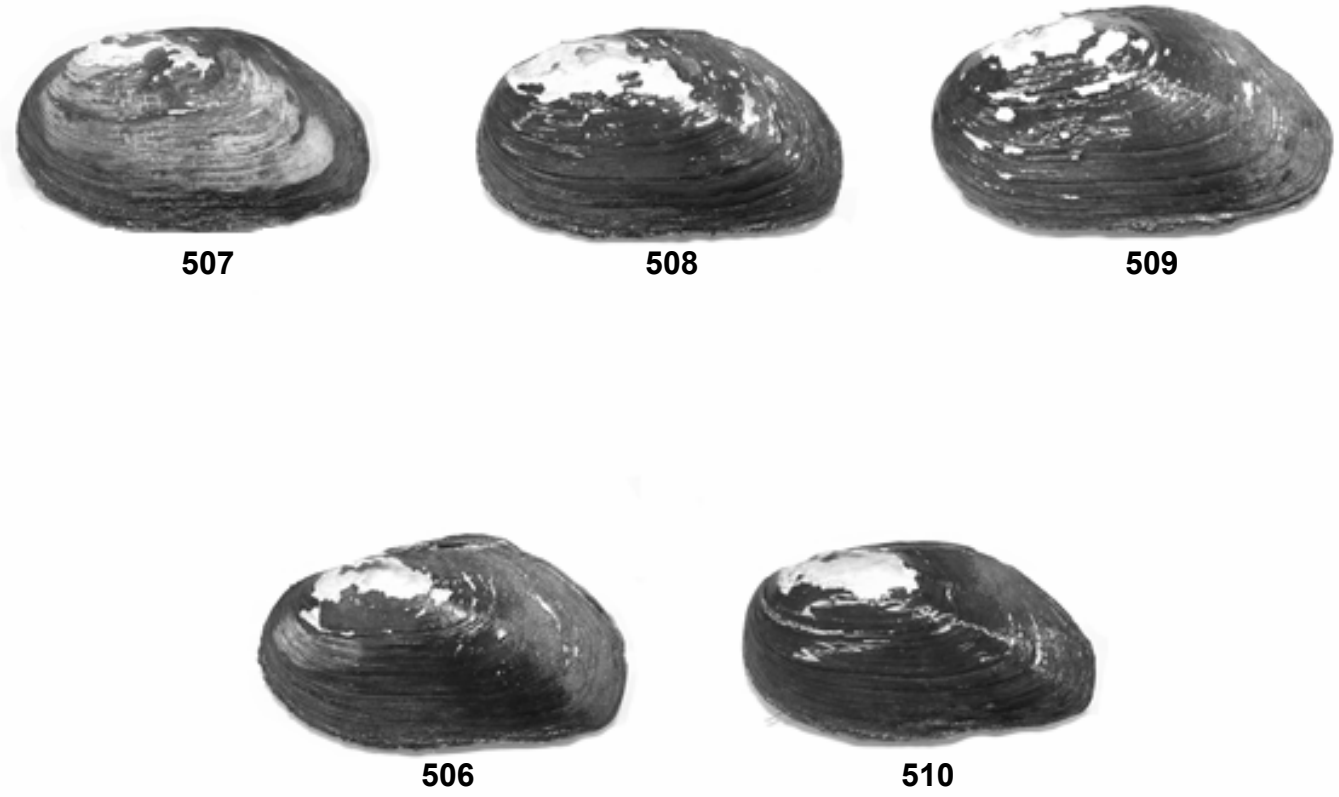


Figure 4.9. Pictures of topotype samples. Lengths of samples are: 506 = 10.1 cm, 507 = 9.83 cm, 508 = 9.58 cm., 509 = 11.2 cm, 510 = 10.6 cm.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

Conclusion

Elliptio complanata is considered the most abundant and widely distributed freshwater mussel native to the Atlantic slope. However, field observations have consistently suggested that there is either marked site-specific morphologic plasticity or that the field specimens we routinely identify as *E. complanata* are actually multiple species. These studies were initiated to provide a preliminary contribution to the on-going and long-term effort of understanding the species complexity of *E. complanata*. Together, the body of work presented provides preliminary documentation that *E. complanata* collected from different locations display marked morphological and genetic variation.

Variation in shell morphology was examined to document morphologic differences of *E. complanta* obtained at seven locations in the Cape Fear River (CFR) and one topotype site, Potomac River, Washington D.C. Thirty shell landmarks were used. Based on these, morphological differences between *E. complanata* from CFR and the topotype sites were shown. Out of the thirty morphologic landmarks, only seven were suggested as possible variables to distinguish the collected *E. complanata* from CFR. The seven variables describe the shells according to measures of shell inflation (or obesity) and thickness of posterior and anterior angles.

Nucleotide and molecular diversity based on the COI gene within *E. complanata* from CFR samples suggested that populations were genetically structured based on river location (chapter two). Differences in nucleotide diversity were observed among individuals

found in upstream vs. downstream sites. Although the molecular diversity of *E. complanata* collected within one creek from the CFR supports Ortmann's theory on shape and station, a broader investigation including individuals from larger rivers and streams still has to be performed. Evidence of COI gene variability was also demonstrated through the existence of four major groups within *E. complanata* from CFR. The four phylogenetic groups based on COI gene sequence, showed lack of correlation with four groups from cluster analysis of morphometry data based on seven shell landmarks. This incongruence implied that the seven shell landmarks were possibly insufficient in distinguishing *E. complanata* from CFR. Furthermore, the incongruence between morphometry and COI sequence data suggested the reliability of genetic data over the morphometric variables for the very plastic *E. complanata* species.

Morphometric and genetic analyses suggested differentiation between *E. complanta* obtained in the CFR and from topotype sites. Once again, morphometric differences between topotypes and CFR samples were demonstrated. Based on F_{ST}, Potomac site (topotype site) was genetically different from the rest of CFR sites. In addition, genetic homogeneity was revealed within topotypes based on the COI gene.

Genomic fingerprinting corroborated and strengthened earlier claims recognizing high variability within *E. complanata* (Davis et al. 1981). Genomic markers demonstrated more variation within topotypic specimens than COI sequences. Genomic variation was apparent from the high heterozygosity values calculated within the topotypes. In addition, cluster analysis of AFLP markers did not cluster all five topotype specimens into one group, and suggested marked genomic variability within topotypes. Phenotypic differences of the topotypes, based on overall shell form, seemed to agree with the genomic variation derived.

Empirical investigations involving more topotype samples need to be performed to support the apparent agreement between morphological and genomic variation observed within topotypic materials.

Genomic data defined more groups than the single group derived from COI sequencing. Specifically, three distinct and consistent groups resulted from cluster analysis of AFLP data, two of which contained topotypic specimens. The dispersal of the topotypes into the two AFLP groups supports earlier the claim that there is high variation within the topotypic specimens.

Despite the logistical constraints encountered during data collection, which led to the small sample sizes used for genetic analyses, the trends presented in this research still suggest the serious need for reconsideration of the species group, *E. complanata* as suggested by earlier researchers (Bogan et al. 2003). A more thorough analysis of genetic and morphological variation among purported *E. complanata* from the eastern U.S. where they occur in great abundance is highly recommended. The use of sequence data from non-mitochondrial genes, paired with a more comprehensive morphological analysis had been recommended for assessing species identity of *Anodonta* in the western U.S. (Mock et al. 2004). This same recommendation should also be performed for *E. complanata* in the eastern U.S. so its species identity and relationships with other freshwater mussels maybe established. In general, the use of a single gene such as the COI gene has been reported to provide incomplete information on the true population history of a species (Ballard and Whitlock 2004). A species' evolutionary history may be better understood if genetic information was derived from more than one gene source, which would provide an independent estimate of the species tree (Ballard and Whitlock 2004). Genomic

fingerprinting, specifically AFLP seems to provide a way to circumvent the shortcoming presented by gene sequencing as markers are generated from the whole genome of an individual, even without prior knowledge of its actual genomic sequence. Initial findings from this study showing genomic variation within the topotype group, which are supposedly monophyletic based on COI sequence data, should be corroborated by the application of AFLP on more *E. complanata* collected from other locations that were historically reported to contain this species. If genomic variation still exists, then we can conclude with greater certainty that techniques such as AFLP provide a more extensive description of genetic variation for *E. complanata* than COI sequencing.

Detailed morphological assessment maybe performed by testing the occurrence of the seven conchological landmarks reported in this study with more individuals from other locales, especially in the southeast U.S. Having specific conchological features to differentiate *E. complanata* would be especially useful to field biologists who depend on shell form and meristics in initially identifying freshwater mussels. In addition, internal anatomical features should also be included as possible additional characteristics that may distinguish the *E. complanata* individuals found at different sites. Combined conchological and internal structures may clearly identify the varying forms of *E. complanata* and may assist in arriving at different actions, which could lead to improved management practices for specific populations.

Recommendations

Further characterization of *E. complanata* is a potentially huge undertaking that could involve years of intensive research. These studies lead to numerous questions and possibilities, which warrant more investigation. Taken together, the results presented in this study are sound preliminary findings that need to be corroborated if we want to clarify species identity and relationships of *E. complanata*. Specific recommendations suggested for the continuation of the study of *E. complanata* species complexity include:

1) Intensive morphological and genetic analyses of *E. complanata* using a larger sample size. In this study, morphology and COI-based correlation did not agree with the shell form and genomic variation observed from AFLP data. The lack of concordance suggests the need for more empirical investigation involving a bigger sample size. For the topotypes, a larger sample size is needed to characterize with greater certainty that the variation observed is a true representation of a typical *E. complanata* and can then be used for comparisons with other individuals found at different sites. In general, a larger sample size could provide higher discriminatory power for the statistical tests. A sample size of more than ten individuals was suggested to provide powerful statistical estimates (Berg and Berg 2000).
2) Development of non-invasive techniques for collection of DNA material. The development and use of a non-invasive technique for genetic analysis will help future researchers avoid the logistical constraint that I encountered in my sample collection. My sample collection was restricted because of the invasive nature of the tissue collection that I have employed, which required sacrificing the collected animals.

Haemolymph-based genetic analysis may hopefully provide a welcome alternative that will be beneficial for the animals, as well as researchers and conservation managers.

3) The application of microsatellite analysis to corroborate genetic diversity results from AFLP. Microsatellites provide specific advantages when contrasted with AFLP markers. They provide a means of determining allelic status, and its hypervariability allows for the detection of fine population structure. Its disadvantage lies in the small number of microsatellite loci produced for genetic analysis, which becomes a drawback when performing statistical analysis. In this regard, markers derived from AFLP provide a larger sample size for statistical analysis and are thus more robust. Combining results from hypervariable codominant microsatellite markers and dominant AFLP markers should prove to be a powerful method for examining population structure of *E. complanata*.

4) Gene sequencing of samples to determine the relationships of purported *E. complanata* with other freshwater mussels. Identifying whether species groups are synonymous with another valid species or whether another valid species group exists, are important in establishing conservation/management measures for *E. complanata* and the genus *Elliptio*, in general. We need to protect the diversity present in the remaining *Elliptio* species, before they suffer the fate of other extinct and endangered *Elliptio* species (O'Brien et al. 2003).

REFERENCE

Ballard, J.W.O. and M.C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729 - 744.

Berg, D. J.and P. H. Berg. 2000. Conservation genetics of freshwater mussels: comments on Mulvey et al. Conservation Biology 14(6): 1920-1923.

Bogan, A.E., J.M. Serb and C. Lydeard. 2003. Determination of the systematic position and relationships of the *Elliptio* complex (Mollusca: Bivalvia: Unionidae) with distributions in North and South Carolina. A Final Report submitted to USGS (Contract 01HQAG0185).

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

Mock, K.E., J.C. Brim-Box, M.P. Miller, M.E. Downing and W.R. Hoeh. 2004. Genetic diversity and divergence among freshwater mussel (*Anodonta*) populations in the Bonneville Basin of Utah. Molecular Ecology.

O'Brien, C.A., J.D. Williams and M.A. Hoggarth. 2003. Morphological variation in glochidia shells of six species of *Elliptio* from Gulf of Mexico and Atlantic Coast drainages in the southeastern United States. Proceedings of the Biological Society of Washington 116(3): 719-731.

APPENDIX A

Protocol for mitochondrial DNA (mtDNA) sequencing of cytochrome oxidase I or COI sequencing

I. PCR amplification:

A. In 1.5 ml microfuge tube, create Master Mix:
(n = 1) 2.5 μl LC 01490 (forward primer)
2.5 μl HC02198 (reverse primer)
20 μl Hotmaster Taq
23 μl sterile distilled H_2O

B. Mix 2 μl DNA template with 48 μl master mix by centrifuge or by pipetting master mix up and down the centrifuge tube.

C. Place solution in thermocycler using the following program:
One cycle: Step1: 94oC for 1 minute (Method 1)

40 cycles: Step 2: 94oC for 1 minute
Step 3: 50oC for 1 minute } (Method 2)
Step 3: 72oC for 2 minute

Step 4: 72oC for 5 minutes (Method 3)

Step 5: 4oC forever (Method 4)

NOTE: I used the Perkin-Elmer thermocycler that is why there was a need to assign different 'Methods'. For this thermocycler, there was a need to assign a method name or number for every step and to link the different methods to run the program. I linked the four programs and saved it under 'Method 5'.

D. Clean PCR product and freeze samples for later steps.

E. Check PCR product quantity and quality on 1% agarose mini-gels with ethidium bromide using 100 bp ladder as standard.

II. PCR clean up and drying

A. Thaw the following from the freezer and store on ice:
cleaned PCR products
BigDye terminator tube
LC01490 tube
HC02198 tube

You may facilitate thawing by centrifuging the tubes for a few seconds and storing them on ice immediately OR you may thaw tubes in fridge for several hours and take out of fridge when ready to use.

B. Label new PCR tubes (0.2 μl) – one for the forward 1o and the other for the reverse. You will need two tubes per sample.

C. Create master mix for each sequencing primer-reaction combination (i.e. one master mix for each primer). Each master mix has the following combination:

(n = 1) Master mix for forward reaction:	BigDye	2.0 μl
	LC01490	0.5 μl
Master mix for reverse reaction:	BigDye	2.0 μl
	HC02198	5 μl

D. Take the new labeled tubes in B and fill each with 2.5 *u*l of template (cleaned PCR products from step I)

E. In each tube in D, put 2.5 *u*l of master mix. Mix solution by centrifugation or pipetting solution up and down each tube.

F. Run reaction in thermocycler following this program:

One cycle: Step 1: 96^{o}C for 3 minutes

30 cycles: Step 2: 96^{o}C for 10 seconds
Step 3: 50^{o}C for 5 seconds
Step 4: 60^{o}C for 4 minutes

Step 5: 4^{o}C forever

III. Reaction clean up and drying

A. Add 15 μl of sterile distilled water to bring volume up to 20 μl.

B. Spin them down using inserts for the centrifuge.

C. Prepare a DyeEx column for each reaction (Qiagen):
   a. Vortex each column to resuspend the gel matrix in the tube
   b. Open the tube slightly, break off the tab on the bottom of the column and place in a collection tube
   c. Centrifuge the column set-up for 3 minutes at 2800 rpm

D. Carefully move the tube with the prepared resin to a new 1.5 ml tube and discard the (now wet) collection tube.

E. Transfer sequencing reactions to the center of the resin bed in the column.

F. Spin the reactions for 3 minutes at 2800 rpm

G. Discard the resin and label the tubes

H. Open the 1.5 ml tubes and place them in the centrivap (make sure you balance the samples to assure efficient operation of equipment).

I. Connect the black tube on the back of the centrivap to the house vacuum and plug it in.

J. Turn on rotor and heater. Let it come to full speed

K. Then, apply the vacuum.

L. Let it run for 1 hour to 1 hour 10 minutes or until reactions are almost dry to dry.

M. Cap them and place in the refrigerator until ready for gel loading.

IV. Gel preparation

A. Prepare the following reagents:
   a. 10% Ammonium Persulfate (APS)
      - Prepare 10% APS no more than 2 hours before pouring the gel

      - Weigh 0.05 +/- 0.005 g of ammonium persulfate into a 15 ml polypropylene tube.
      - With a P-5000 Pipetman or equivalent, add 5 ml of deionized water to the tube.
      - Vortex until all crystals dissolve.
   b. Long Ranger Gel
      - This recipe is for a 36 cm, 5% Long Ranger/6 M urea gel (these are what is typically used for ABI377 sequencer)

      - Weigh out 18.0 g of urea and transfer it carefully to a graduated cylinder.
      - Using a pipette, add the following:
        5.0 ml 50% Long Ranger gel solution concentrate
        5.0 ml 10x TBE
      - Slowly add distilled, deionized H2O to bring the liquid level to ~ 45 ml. Gently tap the cylinder while adding water to release any air bubbles trapped by the urea.
      - Stopper and invert the cylinder to dissolve the urea. Although the cylinder and its contents become very cold, the urea dissolves very rapidly.
      - Allow the solution to warm to room temperature.

- Add distilled, deionized H2O to a final volume of 50 ml.
- Stopper the cylinder and then mix the contents thoroughly.
- Filter the mixture through 0.2 μm cellulose nitrate filter.
- Degas the filtrate by one of the following methods:
  Purging with argon or helium
  Applying a vacuum for at least five minutes
- Proceed immediately to the next step (B)

B. Glass plate preparation
   a. Wash and rinse top and bottom glass plates (size and thickness) with 1% Alconox solution and warm water.
   b. Squirt with 70% ethanol and wipe with wipes (brand?) until squeaky clean and free of dust.
   c. Put clean bottom plate on holder (I used old Styrofoam tube container as holder).
   d. Put slightly wet (with distilled water) side spacers on left and right sides of bottom plate.
   e. Place clean top plate on bottom plate. Check to make sure that notched side of both plates are facing outside.
   f. Fit clamping rails on left and right side of plates. Finger tighten the clamps except the topmost clamp.
   g. Set aside until ready for gel pouring.

C. Gel Polymerization
   a. Add the freshly made 250 μl 10% APS to the 5% Long Ranger gel solution. Swirl gently to mix BUT be careful to avoid introducing air bubbles.
   b. Add 25 μl TEMED (tetramethylethylenediamine) to the solution. Swirl gently to mix BUT be careful to avoid introducing air bubbles.
   c. Immediately cast the gel. Use standard 0.2 mm spacers and a square tooth comb.
   d. Allow the gel to polymerize for at least 2 hours before performing electrophoresis

V. Gel electrophoresis (based on ABI377 protocol)

A. Sample preparation
   a. Resuspend samples into 2.5 μl of ABI loading dye
   b. Vortex briefly.
   c. Spin down briefly.
   d. Keep in fridge or freezer or ice until ready to load.

B. Buffer preparation
   a. Prepare 1400 ml of 1x TBE running buffer from a 10x TBE stock in a graduated cylinder (1260 mls dH2O + 140 mls 10x TBE).

C. Preparation of Gel for electrophoresis
   a. Clean gel with dH2O until squeaky clean and free of dust and fingermarks.

b. Check well and rid it of polymerized gel.
c. Allow to air dry.

d. Restart the computer
e. Open "ABI Prism 377 – 96 Collection"
f. Open "File", click "New", select "Sequence sample"
g. Create sample sheet
    i. Check dye/set primer
    ii. Type in sample names
    iii. Save
h. Go to "File", click "New", select "Sequence run"
i. Check the number of lanes (i.e. 36, 48 or 96)
j. Select sample sheet
k. Check run module: Seq. run 36E-2400copy
l. Check pre run module: Seqpr36E-2400
m. Wipe off gel with distilled H2O
n. Place in frame
o. Place comb
p. Place gel in sequencer
q. Start plate check
r. If OK, cancel plate check
s. Put on upper buffer tank
t. Place heat plate
u. Add buffer (from VB)
v. Place lid on upper tank, then close door
w. Start pre-run, go to "Window", select "Status", check settings, continue pre-run until temperature get around 52 degrees
x. Pause pre-run, blow out wells with plastic syringe
y. Load samples starting with the odd numbers, start at #7, finish with #'s 5, 3, and 1. Replace lid and resume pre-run for 5 minutes
z. Go to "Window", select "Status" and watch time.
aa. Pause pre-run, blow out well with syringe
bb. Add even # samples starting with #8 and finishing with #'s 6, 4 and 2.
cc. Replace lid and close door.
dd. Cancel pre-run, when asked to terminate click "Yes"
ee. Start run, save as "Gelfile(date)". Go to "Window", select "Status" and check settings.
ff. Machine will run for 4 hours and will automatically shut off when run is finished.

VI. Data compilation of sequence:

A. From ABI file of colored map, trace files were produced. In ABI program trace files were read and imported to the program, Sequencher.

B. In the program, Sequencher, trace files were opened where the forward and reverse sequences were matched, aligned, compiled and proofread (sequence files were considered 'clean' at this point).

C. 'Cleaned' sequence files were then opened in PAUP*4.0 into a nexus file format (format that is default for PAUP).

D. Neighbor joining and maximum parsimony analyses were then performed in PAUP*4.0.

E. From PAUP*4.0, nexus formatted data were exported to Phylip file format so it could be opened in the program Collapse (1.1). This program allows the dataset to be collapsed into haplotypes which makes creating and reading trees easier. Collapse is a freeware by David Posada (http://inbio.byu.edu/Faculty/kac/crandall_lab/programs.htm)

APPENDIX B

AFLP Protocol

(Adapted from LICOR manual and Myburg, A.A. and D.L. Remington. 2000. Protocol for high-throughput AFLP analysis using LI-COR *IR*^{2} Automated Sequencers).

NOTE: Roman numeral assignments were those that were used as labels on tubes

All steps should be performed in ice! Master mixes should be prepared in properly labeled 1.5 ml centrifuge tubes. Before performing the first step (I), label individual 0.2 ml PCR tubes or 96-well tubes.

I. Restriction of DNA template (Time: > 2 hrs 15 mins)

A. Purchase LI-COR's AFLP Template Preparation Kit. On ice and in 0.2 ml PCR tube, create master mix containing the following:

(n = 1) 5X reaction buffer	2.5	μl
Template DNA (100 ng in <= 9 μl)	<= 9.0	μl
EcoR1/Mse1 enzyme mix	1.0	μl
Deionized water	to 12.5	μl
Total volume	12.5	μl

B. If using individual PCR tubes, cap them; if using 96-well plate, put an appropriate seal on top. Centrifuge briefly and put mixture in thermocycler using the following program:

One cycle: Step 1 37^{o}C for 2 hrs
Step 2 70^{o}C for 15 minutes
Step 3 4^{o}C forever

II. Ligation (Time: 2 hrs)

A. From the AFLP Template Preparation Kit, create master mix (on ice):

(n = 1)	Adapter mix	12.0	μl
	T4 DNA ligase	0.5	μl
	Total volume	12.5	μl

B. Mix solution by centrifugation or by pipetting mixture up and down the tube.

C. Remove the tubes in thermocycler and add 12.5 μl of master mix into each tube. This gives you a final volume of 25 μl (restriction solution from I + master mix of IIA) of ligation mixture.

D. Mix solution gently and centrifuge briefly. Put back tubes in thermocycler with the following setting:

One cycle: Step 1 20^{o}C for 2 hours
Step 2 4^{o}C forever

E. Dilute ligation mixture (1:10) by transferring 10 μl of the mixture to a new 0.5 ml centrifuge tube and adding 90 μl of TE buffer. Mix well.

F. Store unused portion of ligation mixture (15 μl) at –20oC for long-term use.

III. Pre-amplification (Time: 15 to 20 minutes)

A. Transfer 2.5 μl of the diluted ligation mixture (from previous step IIE) to a new and properly labeled 0.2 ml PCR tube. Store in ice. Store unused portion in –20^{o}C for long-term use.

B. Create master mix using the following:

(n = 1)	AFLP Pre-amp primer mix	20.0	μl
	PCR 10x reaction buffer*	2.5	μl
	Taq DNA polymerase	0.5	μl
	Total volume	22.5	μl

* I used QIAGEN's Taq polymerase kit

C. Add 22.5 μl of master mix into each 0.2 ml PCR tube containing 2.5 μl diluted ligation mixture. This gives you a total volume of 25.0 μl pre-amplification mixture.

D. Cap tightly and centrifuge or pipette mixture up and down each tube and seal with 96-well seal.

E. Put tubes in thermocycler following this program:

20 cycles: Step 1 94^{o}C for 30 seconds
Step 2 56^{o}C for 1 minute
Step 3 72^{o}C for 1 minute

Step 4 4^{o}C forever

F. Perform a 1:40 dilution by pipetting 5 μl of the pre-amplification DNA into a 0.5 ml microcentrifuge tube and adding 195 μl of ddH2O. Store tubes on ice.

G. Store unused portion of F in –20oC for long-term use.

IV. Final amplification (Time: 1 hour)

A. Create master mix in in 1.5 ml microfuge tube (adjusted depending on the number of samples to be run) IN ICE and MINIMIZE LIGHT EXPOSURE! The master mix contains the following:

(n = 1) 10x PCR buffer	2	μl
dNTP mix	1.6	μl
*Eco*RI primer	0.83	μl
*Mse*I primer	5	μl
Taq polymerase	0.24	μl
Distilled H2O	5.33	μl
Total volume	15.0	μl

B. In 0.2 μl PCR tube, take 15 μl of master mix solution and add it to 5 μl of diluted pre-amp mixture (step III F). Again, this should be done IN ICE and IN MINIMAL LIGHT!

C. Mix the new solution – if you are using a 96-plate well and you don't have a centrifuge that can accommodate such a plate, you can mix by pipetting each sample up and down the each tube; otherwise use the centrifuge with plate attachment and mix for a few minutes. If you are using individual PCR tubes (just as I was), mix the solution in our centrifuge that has attachments for 0.2 μl tubes.

D. Close tubes (if using plates and MJ thermocycler, use MJ seals; if using tubes, individually cap them).

E. Run in thermocycler. I used the MJ thermocycler (at C-322) and saved the program "FINAL" which has the following configurations:

13 cycles:	Step 1 94°C for 10 sec	Denaturation
	Step 2 65°C for 30 sec, less 0.7° per cycle after the first cycle	Annealing
	Step 3 72°C for 60 sec	Extension
	Step 4 go back to Step 1 = 13 x	
	(For MJ thermocycler, utilize the 'increment' procedure)	
25 cycles:	Step 5 94°C for 10 sec	Denaturation
	Step 6 56°C for 30 sec	Annealing
	Step 7 72°C for 60 sec, plus 1 sec per cycle	Extension

Step 8 go back to Step 5
(For MJ thermocycler, utilize the 'extend' procedure)

1 cycle: Step 9 72°C for 2 minFinal extension

Hold: Step 10 4°C forever

Step 11 End

F. Transfer samples in properly labeled 1.5 ml tubes. If your lab is equipped with a speedvac, skip this step and proceed to the next.

G. Dry samples in speedvac for approximately 1 hr and 5 mins or until samples are just about dry at the bottom of the tube. Minimize light exposure by covering the top of speedvac with aluminum foil.

H. Add 8 μl of blue stop solution.

I. Denature for 3 mins using an incubator at 90°C for 3 minutes.

J. On ice, transfer samples into 96-well plate. Make sure samples follow the correct orientation on the plate. For my run, I have 62 samples to be loaded on a 64 lane gel. I used the following orientation:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	*1*	*2*	*3*	*4*	*33*	*34*	*35*	*36*	*65*	*66*	*67*	*68*
B	*5*	*6*	*7*	*8*	*37*	*38*	*39*	*40*	*69*	*70*	*71*	*72*
C	*9*	*10*	*11*	*12*	*41*	*42*	*43*	*44*	*73*	*74*	*75*	*76*
D	*13*	*14*	*15*	*16*	*45*	*46*	*47*	*48*	*77*	*78*	*79*	*80*
E	*17*	*18*	*19*	*20*	*49*	*50*	*51*	*52*	*81*	*82*	*83*	*84*
F	*21*	*22*	*23*	*24*	*53*	*54*	*55*	*56*	*85*	*86*	*87*	*88*
G	*25*	*26*	*27*	*28*	*57*	*58*	*59*	*60*	*89*	*90*	*91*	*92*
H	*29*	*30*	*31*	*32*	*61*	*62*	*63*	*64*	*93*	*94*	*95*	*96*

For a 48-well lane, the following arrangement should be followed:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
A	*1*	*2*	*3*	*25*	*26*	*27*	*49*	*50*	*51*	*73*	*74*	*75*
B	*4*	*5*	*6*	*28*	*29*	*30*	*52*	*53*	*54*	*76*	*77*	*78*
C	*7*	*8*	*9*	*31*	*32*	*33*	*55*	*56*	*57*	*79*	*80*	*81*
D	*10*	*11*	*12*	*34*	*35*	*36*	*58*	*59*	*60*	*82*	*83*	*84*
E	*13*	*14*	*15*	*37*	*38*	*39*	*61*	*62*	*63*	*85*	*86*	*87*
F	*16*	*17*	*18*	*40*	*41*	*42*	*64*	*65*	*66*	*88*	*89*	*90*
G	*19*	*20*	*21*	*43*	*44*	*45*	*67*	*68*	*69*	*91*	*92*	*93*
H	*22*	*23*	*24*	*46*	*47*	*48*	*70*	*71*	*72*	*94*	*95*	*96*

K. Let samples stay on ice or in –20^{o}C freezer, until ready for gel loading.

V. Gel preparation and casting

A. Glass plate preparation

a. Wash and rinse top and bottom glass plates (25 cm and 0.25 mm) with 1% Alconox solution and warm water.
b. Squirt with 70% ethanol and wipe with Kimwipes until squeaky clean and free of dust.
c. Put clean bottom plate onto old but clean Styrofoam tube container (or you may also use two plastic pipet tip box covers as holder.
d. Put slightly wet (with distilled water) side spacers (thickness: 0.25 mm) on left and right sides of bottom plate
e. Place clean top plate on bottom plate. Check to make sure that notched side of both plates are facing outside
f. Fit clamping rails on left and right side of plates. Finger tighten the clamps except the topmost clamp
g. Set aside until ready for gel pouring.

B. Gel preparation

h. Prepare the following solution:
   - 0.8x TBE = 80 ml of 10x TBE + 920 ml of distilled H2O
   - 8% Long ranger stock solution*
   - 10 % ammonium persulfate (APS)
   - TEMED

*8 % Long ranger stock solution:

100 ml	42.04 g	Urea (7M0
	15 ml	50% Long Ranger
	6.4 ml	10x TBE
	to 100 ml	dd H2O

i. Mix 20 ml Long ranger solution, 150 □l of 10% APS and 15 □l of TEMED in a small squirt bottle
j. **Gently** swirl the resulting solution for a few seconds.

C. Casting the gel

k. Quickly squeeze or squirt the gel solution into the notched space of prepared glass plate. Squirt all the contents from the depressed top portion of the glass plates. When you squirt, move the bottle continuously from side to side and DON'T stop pressing the bottle, lest you develop bubbles inside the plate. The gel solution will flow between the plates by capillary action.
l. In case bubbles form, you may remove it by inserting a radiography wire in between plates and pull bubble/s out. Tapping lightly the glass plates may aid

in the flow of solution towards the bottom part of the plates. Let solution overflow to the bottom end of the plates.

m. Then, quickly but gently insert flat side of 48- or 64- well comb (thickness: 0.25 mm) on the center. Make sure that the thickness of the comb is the same as the thickness of the spacers (*i.e.* if you used 0.25 cm spacers, then use the same comb thickness) and that you don't introduce bubbles.
n. Put gel-casting plate and finger-tighten the uppermost screws of clamping rails.
o. Put plastic cling wrap around glass plates and let stand for at least 90 minutes at room temperature. You may leave the glass plates (with unhardened gel) in the fridge for several hours. You may also prepare this step the night before and leave the glass plates (with unhardened gel) in the fridge – make sure you run the gel immediately first thing in the morning, otherwise the gel would continue to degrade.

VI. Running the gel

A. Assembly of gel onto LI-COR4200 sequencer
   a. If the gel has already hardened, rinse the glass plate with warm, soapy water (1% alconox). Use a soft scrub or sponge in removing the polymerized gel outside the plates – pay particular attention in cleaning the notched and bottom part of plate because that is where the machine's infrared detector will pass.
   b. Remove the comb on the notched part of the plate. Rinse this portion also.
   c. Dry the plates by wiping it off with 70% alcohol and lint-free Kimwipes until it is free of dust and fingermarks, especially the notched and bottom part of the plate. Be liberal in using alcohol-saturated Kimwipes.
   d. Dry the notched part of the plate by inserting a dry Kimwipe into it. Make sure that area is free of gel bits and that it is dry
   e. Dry the clamp rails and the outside of the plates with Kimwipe to avoid short circuit or power leakage which will lead to sequencer shut down.
   f. Hold the plates up against the light for inspection. Make sure the bottom part of the gel where the laser will pass is free from dirt, dust and prints so the image quality of the gel will not be affected.

B. Setup of automated sequencer
   a. Prepare 0.8x TBE as LI-COR running buffer:
      80 ml 10x TBE
      920 ml distilled water
   b. Place the lower buffer tank against the base of the sequencer.
   c. Place the gel plate sandwich upright into a gel casting stand.
   d. Dampen a 48- or 64-well comb with distilled water or TBE and insert it through the plates, completely centered relative to the black heating plate of the sequencer. Make sure that only the tips of the teeth are inserted and once comb is inserted, DO NOT REMOVE IT OR ATTEMPT TO MOVE IT OUT OF THE PLATE. This might lead to sample leakage and/or breaking of gel.

e. Hang a silver reflector screen along the upper backside of the backplate. This helps improve the visibility of the top of the gel.
f. Hold the gel assembly firmly by both clamp rails, slide the plate sandwich into the slots along the side of the heater plates of the sequencer.
g. Secure a white gasket into the gasket groove of the top buffer tank. Dampening it with distilled water would help in securing it onto the groove.
h. Completely unscrew the top clamp knobs of the clamping rails attached to the glass plates and insert the top buffer tank. Finger-tighten the top clamp knobs.
i. Fill the top and lower buffer tanks with 0.8x TBE until the maximum fill line.
j. By using plastic transfer pipette, squeeze a few TBE onto the well area to blow away any urea, pieces of gel or remove bubbles.
k. Insert cover of the top and lower buffer tanks. Attach the safety power cord of the top buffer tank, to the buffer tank and on the body of the sequencer.
l. Close the sequencer door and make sure the interlock snaps shut.

C. Setup of the data collection software (e-Seq). Turning on the machine is dependent on starting the software. This portion starts the pre-run of the machine. You will set-up the program and assign a file directory and name.

VII. Sample preparation and loading

A. Sample denaturation
   a. Thaw frozen 96-well plate containing samples. Denature the samples using the MJ thermocycler, following this program:

      One cycle: Step 1 90oC for 3 minutes
      Step 2 4oC forever
      End

   b. Remove LI-COR molecular weight standard (50 – 700 bp) from the freezer and place in an ice bucket to allow it to thaw.
   c. Remove plate from thermocycler and keep on ice until ready for loading.
   d. Cover ice bucket with foil to minimize light exposure.

B. Loading samples
   a. Pause the LI-COR machine by clicking on the "Load samples" option on the program (ON THE COMPUTER SCREEN).
   b. Disconnect the safety power cord of the top buffer tank to the body of the sequencer.
   c. Remove the cover of the top buffer tank.
   d. Using a plastic transfer pipette, clean the wells by squeezing some of the TBE in the tank to the well.
   e. Using an 8-channel Hamilton syringe, practice loading samples by loading dH_2O and expelling 0.7 to 1.0 □l to a pipette tip box.

f. Once ready, load 0.7 to 1.0 □l of each sample, one column of the plate at a time and placing the first sample in the lane marked "1" on the comb. This will load every fourth lane (on 64-well combs) or every sixth land (on 96-well combs). The samples will be in order if the plates were set up as described above IVJ. Amounts loaded will be more uniform if the syringe barrels are completely filled with sample, and then all but the desired loading volume is expelled back into the plate.
g. Rinse the syringes by drawing and expelling several volumes of dH_2O from a container such as a pipette tip box lid, in between samples.

C. Loading standards

When the gel is already loaded with samples, use a single lane Hamilton syringe or one of the end channels of the 8-channel syringe to load ~0.7 □l of the molecular weight standard. Load the standard in the lanes immediately to the left and right of the first and last sample lanes, respectively. You may also load the standard in the center of the gel (lane 33). These assure you that one or two standards will appear on the gel even if the comb is not centered.

VIII. Running the gel

A. When sample loading is complete, replace the lid of the upper buffer tank and the electrical connector.
B. Close the sequencer door.
C. Click on "RUN GEL" icon to start the program.
D. View gel image from the computer screen.
E. The program will automatically end if the run is complete.
F. You may view gel run through this website (www:/152.1.33.32).

IX. Resulting image file (in tiff format) maybe scored and analysed using the program, BIONUMERICS.