2010-08-06

Phylogeography of the Marsh Rice Rat (Oryzomys palustris) in Wetlands of the Southeastern United States

Jane L. Indorf
University of Miami, jane@bio.miami.edu

UNIVERSITY OF MIAMI

PHYLOGEOGRAPHY OF THE MARSH RICE RAT (ORYZOMYS PALUSTRIS) IN WETLANDS OF THE SOUTHEASTERN UNITED STATES

By

Jane L. Indorf

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2010

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

PHYLOGEOGRAPHY OF THE MARSH RICE RAT (ORYZOMYS PALUSTRIS) IN WETLANDS OF THE SOUTHEASTERN UNITED STATES

Jane L. Indorf

Approved:

_____	_____
Michael S. Gaines, Ph.D.	Terri A. Scandura, Ph.D.
Professor of Biology	Dean of the Graduate School
_____	_____
Douglas Crawford, Ph.D.	Barbara Whitlock, Ph.D.
Professor of Biology	Associate Professor of Biology

_____
Wm. David Webster, Ph.D.
Professor of Biology
University of North Carolina Wilmington

INDORF, JANE LEAH (Ph.D., Biology)

Phylogeography of the Marsh Rice Rat (*Oryzomys palustris*) in Wetlands of the Southeastern United States (August 2010)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Michael S. Gaines, Ph.D.
No. of pages in text. (208)

The marsh rice rat (*Oryzomys palustris*) is a semi-aquatic rodent endemic to the southeastern United States. Unlike most terrestrial small mammals, the marsh rice rat can easily disperse over water and has a close association with wetlands. These specialized traits have likely greatly shaped the genetic structure and diversity within this species. I studied genetic patterns within the marsh rice rat to understand how this species' specialized ecology, as well as the geologic and climatic history of the southeastern United States, affected the genetic structuring within this species. The phylogeography of many species in the southeastern United States has been studied and concordant geographic patterns of genetic variation exist among many of these species. Researchers have hypothesized that the biogeography of the southeastern United States has been influenced by the Pleistocene glacial cycles, producing similar genetic patterns within unrelated species. I first examined genetic patterns within the marsh rice rat at the macro scale of phylogenetics. This nominal species actually represents two cryptic species; populations in the eastern and western regions of its range are genetically divergent. I also identified three subspecies, in contrast to the six morphological subspecies historically recognized. The silver rice rat in the Lower Florida Keys and the

Sanibel Island rice rat from Sanibel Island Florida are both subspecific taxa. Only one mainland marsh rice rat subspecies exists. I then studied the phylogeographic patterns within the marsh rice rat and determined that geographic patterns of genetic variation in this species are not concordant with the phylogeographic patterns uncovered in most other species of the southeastern United States. The genetic structuring within the marsh rice rat has been influenced not only by the geologic and climatic history of this region, but also by the species' semi-aquatic adaptation. I also studied genetic patterns at a micro scale by estimating present levels of gene flow and genetic diversity within populations. Gene flow is a contemporary factor in maintaining levels of genetic diversity within populations of the marsh rice rat. From the macro scale of phylogenetics to the micro scale of population genetics, the genetic structure of the marsh rice rat has been shaped by past climatic history and by this species' specialized ecology.

I dedicate this dissertation to

Noel Whitby
My Poppy

Who instilled in me a passion for conservation and ecology, and helped develop my inquisitive nature which has led me down this scientific path.

ACKNOWLEDGEMENTS

First, I thank my advisor, Dr. Mike Gaines, for all his support and guidance over the years. He helped with this research project from beginning to end, from forming my initial research ideas and questions, to preparing my research for public presentation. Mike gave me field training and lent me field equipment necessary to collect samples. He was a wonderful editor on everything from my research proposal, to grant applications, to my final dissertation. His mentorship has been critical to every aspect of my graduate career.

I also thank my committee members, Dr. Doug Crawford, Dr. David Webster, and Dr. Barbara Whitlock. They all imparted great advice, knowledge, and help over the course of my dissertation research. Doug helped me to think about my research in new and different ways, and taught me much about molecular evolution. Barbara gave me additional experience with molecular techniques and from her I learned much about phylogenetics. David trained me in morphometrics, lent me samples, and let me stay with him and his family while visiting his lab at the University of North Carolina Wilmington. I am grateful for his wonderful hospitality.

I thank Dr. Carla Hurt for all her help with my lab work and data analyses. She was always there for me, to help solve problems and give me support when things weren't going well, and when things were going well. I learned much about population genetics from Carla. I also thank Dr. Matt Osentoski, Dr. Daniel Wang, and Dr. Dean Williams for helping me get my project up and running.

I thank the Gaines Lab members, Sean Beckmann, Natalia Borrego, Nathan Dappen, Dr. Ana Ibarra, Dr. Erin Kuprewicz, Janne Nielsen, and Dr. Tiffany Plantan.

Their help, support, and friendship over the years greatly enhanced my time at the University of Miami. I also thank all the other Biograds for their support and friendship.

For help with fieldwork I thank Sean Beckmann, Nathan Dappan, Dr. Mike Gaines, Janne Nielsen, Jason O'Conner, Dr. Tiffany Plantan, and Robert Prendiville. A special thanks to Dr. John Hanson for his research collaboration, for sharing samples, and giving me advice and help with my research. I also thank the Angelo State Natural History Museum, the Museum of Southwestern Biology, and the Museum of Texas Tech University for sample loans. I thank Amanda Crouse, Dr. Kent Edmonds, Phillip Frank, Magaly Massanet, Dr. Bob Rose, Dr. Richard Stevens, and Emily Woods for help with obtaining samples. I also thank John Aspiolea, David Dell, Jean Huffman, Phillip Hughes, Paul Miller, Annette Nielsen, Mark Salvato, and Jenna Wanat for their help in coordinating my fieldwork. I also thank the Florida Fish and Wildlife Conservation Commission for help with permitting. Many others played a role in my successful field collections and I thank them all.

For funding I thank the American Society of Mammalogists, the University of Miami Department of Biology, the University of Miami Department of Biology's Kushlan and Savage Awards, the National Science Foundation and University of Miami's GK-12 Science Made Sensible program, and Dr. Kathryn Tosney.

I thank Robert Prendiville for all his love and support throughout the final years of my dissertation research, especially the final stretch. He kept my spirits up and encouraged me through the writing process. Robert believed in me and in my ability to finish my dissertation.

And finally I thank my family, Anne, Helen, and Gerry Indorf. You helped pull me through the hard times and kept me going when things got rough. Your faith in me is unwavering. I thank you for all your financial support. I am grateful to my mother Anne who flew down to Miami to do my laundry and keep me company while I was writing. I am blessed to have a father who stayed up with me until 2 am, the eve of my defense, to help me practice my presentation. I thank both my parents for all their love. They greatly helped me in accomplishing this goal and gave me the resources and support I needed to pursue a Ph.D.

TABLE OF CONTENTS

LIST OF FIGURES

Page

CHAPTER 2

CHAPTER 3

LIST OF TABLES

Page

Chapter One

Introduction
Phylogeography: The Genetic Signatures of Evolution

Identifying and classifying species is an organic human interest and has been a basic branch of biology since humans began systematically studying the natural world. Phylogenetics, identifying species and determining the evolutionary relationships among them, forms foundational knowledge of organisms. This phylogenetic information is essential for studying a species' ecology, behavior, and evolution. Patterns among species' morphology, behavior, or ecology often are well explained by an evolutionary hypothesis. In order to support such a hypothesis, phylogenetic relationships among taxa must be quantified and patterns of genetic diversity uncovered.

Not only do biologists need to know the evolutionary relationships among species, but also how species evolved and phylogenetic relationships arose. Speciation, the formation of species by divergent evolution, is a complex process and can be driven by a number of different factors. In sexually reproducing organisms, populations may genetically diverge into new species by natural selection, geographic isolation (allopatric speciation: Mayr 1963), the founder effect (peripatric speciation; Mayr 1982), or ecological adaptation with selection (ecological speciation; Rundle and Nosil 2005) among others. Studying these evolutionary processes and their source leads biologists to a deeper understanding of a species' natural history.

The source of evolutionary relationships can be uncovered by studying a species' phylogeography. The analysis of a species' phylogeography can elucidate the evolutionary processes that have shaped a species' genetic differentiation over time

(Avise 2000). Phylogeography evaluates the geographic patterns of genetic diversity within species to infer how past and present geologic and climatic factors have shaped them. Phylogeographic studies also can identify populations or subspecies that may be undergoing speciation. The ultimate source of all genetic diversity is mutation, though much genetic variation within populations also arises through genetic recombination (Futuyma 1998). Population differentiation also is a source for new diversity and may eventually lead to speciation. The evolutionary processes that act on populations maintain biological diversity and produced the vast amount of biodiversity that exists on Earth today.

The field of phylogeography is relatively new and seeks to connect genetic patterns within species to past and present geology and climate in order to better understand biogeography, the distribution of organisms in space and time (Avise et al. 1987; Hickerson et al. 2010). Originally only mitochondrial DNA (mtDNA) markers were utilized for this purpose, but today nuclear gene sequence data and microsatellite loci are routinely employed in phylogeographic studies (Hare 2001). Animal mtDNA is an ideal molecular marker for phylogeographic analyses because it is inherited maternally, has no recombination, evolves rapidly, has a simple genetic structure, and shows extensive intraspecific polymorphisms (Avise et al. 1987). However, mtDNA is only one line of evidence for phylogeographic structuring. Using only a single gene region or marker to identify the genetic relationships among organisms can be problematic. Different genetic markers can have different evolutionary histories due to different effective population sizes, selection, and incomplete lineage sorting (Ballard and Whitlock 2004, Degnan and Rosenberg 2009). Gene trees portraying genetic

relationships among taxa estimated from a single genetic marker can have a different topology from the true species tree (Nichols 2001).

Nuclear DNA markers, which are inherited bi-parentally, may show different patterns of genetic relationships than mtDNA. Nuclear DNA markers generally evolve more slowly than mtDNA and have a larger effective population size, thus nuclear DNA markers could show less genetic structuring than the more rapidly evolving mtDNA (Moore 1995). Also nuclear DNA is subject to recombination, which mtDNA does not usually undergo during meiosis (Birky 2001). Therefore, mtDNA markers allow for more recent intraspecific evolutionary processes to be revealed. Nuclear DNA sequences typically are used to study genetic relationships at the generic level, though the relatively quickly evolving nuclear microsatellite markers can reveal present day population genetic processes such as gene flow (Manel et al. 2003).

Avise et al. (1979a and 1979b) published the first explicitly phylogeographic studies using mtDNA polymorphisms in the white-footed mouse (*Peromyscus leucopus*) and the southeastern pocket gopher (*Geomys pinetis*) in the southeastern United States. Phylogeographic studies have revealed much about the speciation process and biogeography. These studies have addressed the long-standing problem of "What is a species?" and have helped to distinguish individual species (Sites and Marshall 2003, Hickerson et al. 2010). Phylogeography has been used to identify geographic isolation of populations and gain insight into allopatric speciation (Hickerson and Meyer 2008, Krystufek et al. 2009b). Recent population and range expansions, population bottlenecks, and reproductive isolation also have been uncovered through phylogeographic research (Riddle et al. 2008, Hickerson et al. 2010, Schneider et al.

2010). Life history traits, such as mating systems and dispersal, have been studied in a phylogeographic context (Friesen et al. 2007, Gauffre et al. 2009, Estes-Zumpf et al. 2010). Comparisons of phylogeographic patterns within different species inhabiting the same region can be used to infer the effects of past climatic and geologic changes, such as the Pleistocene glaciations, on organisms (Riddle 1996, Bermingham and Moritz 1998, Kholodova 2009). The biogeographic histories of many regions around the globe have been inferred from multi-species phylogeographic studies, such as the Australian Wet Tropics, Europe, and the southeastern United States (Avise 1992, Moritz and Faith 1998, Schneider et al. 1998, Taberlet et al. 1998, Soltis et al. 2006, Carstens and Richards 2007, Weiss and Ferrand 2007). The field of phylogeography bridges many disciplines and has fostered connections between previously isolated fields of research, such as biogeography, paleontology, population genetics, and evolutionary biology (Avise 2009).

Phylogeographic Patterns within Mammals

Since its inception, the field of phylogeography has grown exponentially with mammals being the most popular taxonomic group for phylogeographic research (Beheregaray 2008). A variety of phylogeographic patterns have been found within mammals, starting with Avise's inaugural studies of rodent species from the southeastern United States. These genetic patterns are shaped by species' evolutionary and life histories, as well as regional climatic and geologic history. Large mammal species are generally more vagile than smaller mammals; therefore large and small mammals may show different phylogeographic structuring. For example, little phylogeographic structuring is expected to be present in highly vagile large mammals over a larger

geographic area, and over a smaller geographic area in small mammals. Species with short dispersal distances are more likely than those with large dispersal distances to show phylogeographic breaks caused by geographic barriers (Bond et al. 2001). However, species with short dispersal distances can also show breaks that are not a result of geographic barriers (Irwin 2002).

Even within highly vagile large mammal species different phylogeographic patterns have been uncovered at different geographic scales. The African savannah elephant (*Loxodonta africana*) is capable of migrating over long distances and can survive in many different habitat types, from dry desert to mesic forest (Nyakaana et al. 2002). Despite this ability, deep phylogeographic structuring within this species exists across Africa (Eggert et al. 2002). Comstock et al. (2002) found little to no gene flow across the African continent, between north-central populations and eastern and southern populations. However, populations within eastern and southern Africa showed little phylogeographic structuring indicating the presence of gene flow among populations in these geographic regions. Large scale phylogeographic structuring within this species may have been caused by Pleistocene climate changes (Nyakaana et al. 2002).

Some highly vagile large mammals exhibit little to no phylogeographic structuring, such as the red kangaroo (*Macropus rufus*), which inhabits the arid and semi-arid regions of Australia (Clegg et al. 1998). The red kangaroo is capable of long distance dispersal over 100 km (Croft 1991). Long-range movement of this species is highly dependant on environmental conditions such as drought. Based on mtDNA sequence data, Clegg et al. (1998) proposed that genetic connectivity over this species' range has been maintained for a long period of time. However, both the western grey

kangaroo (*M. fuliginosus*; Neaves et al. 2009), which is distributed across southern Australia, and the eastern grey kangaroo (*M. giganteus*: Zenger et al. 2003), ranging along the east coast of the continent, showed distinct phylogeographic structuring across their entire ranges. Neaves et al. (2009) attributed this genetic structuring to regional differences in habitat preference, among other factors such as historical climate changes. On a smaller geographic scale (< 230 km), the eastern grey kangaroo showed weak genetic structuring indicating high levels of gene flow among populations (Zenger et al. 2003), similar to the phylogeographic pattern seen within eastern and southern populations of the African savannah elephant.

In Rocky Mountain bighorn sheep (*Ovis canadensis canadensis*), highly divergent mtDNA haplotypes were found within herds, as well as many geographically widespread haplotypes among herds (Luikart and Allendorf 1996). The authors attributed this lack of phylogeographic structuring to the absence of population isolation by past geographic or environmental barriers, as well as the presence of gene flow on a regional scale in the recent past. Bighorn sheep range from southwestern Canada to Mexico (Forbes and Hogg 1999), but this study only examined populations of one subspecies in Alberta, British Columbia, Montana, Idaho, Wyoming, and Colorado. As the above examples of African elephants and kangaroos illustrate, phylogeographic patterns vary depending on the geographic scale of the study. More distinct phylogeographic structuring within bighorn sheep is likely to exist across the species' entire range.

Ramey (1995) studied the phylogeography of bighorn sheep in the southwestern United States and Mexico. For comparison the author included some samples from the Rocky Mountain subspecies. The author found no shared mtDNA haplotypes between

the two regions, indicating that there may be phylogeographic structuring between populations in the two different geographic regions. In a study based on microsatellite loci, Forbes and Hogg (1999) concluded that there is genetic differentiation among populations of the Rocky Mountain bighorn subspecies. The use of different markers among the different studies produced different results; the mtDNA markers may have detected past genetic history, while the microsatellite markers uncovered contemporary genetic relationships among populations.

Moose (*Alces alces*) show the phylogeographic structuring typically expected of large vagile mammals (Cronin 1992, Hundertmark et al. 2002). Within regions and across the species' entire range no phylogeographic structuring was detected. The authors attributed this to a recent population expansion following the last glacial maximum, as well as gene flow (Hundertmark et al. 2003). The same pattern has been uncovered in the coyote (*Canis latrans*) and the gray wolf (*C. lupus*) in North America (Vilà et al. 1999). As with the moose, the author attributed this lack of phylogeographic structuring to population expansion following the last glacial period.

Phylogeographic patterns in small mammals also are often attributed to the effects of Pleistocene glaciations. Unlike large mammals, which seem to show a spectrum of phylogeographic patterns, small mammals usually exhibit deep genealogical subdivisions across geographic regions (Avise 2000). This pattern is found in many North American small mammal species. Phylogeographic patterns are not always simply explained by past climatic events, but often a combination of past and present effects may have shaped the genetic structure of a species.

The American red squirrel (*Tamiasciurus hudsonicus*), which is one of the most common arboreal species inhabiting subalpine and montane forests in North America, is dependent on boreal vegetation (Steele 1998). Because this habitat requirement generates a fragmented distribution, the red squirrel most likely has deep phylogeographic structuring across its range. In the central Rocky Mountain region, the red squirrel is divided into two highly divergent clades: one in southern Colorado and the other in northern Colorado, Wyoming, eastern Utah, and eastern Idaho (Wilson et al. 2005). The authors attributed this pattern to the Green River in the Wyoming Basin. The river and surrounding habitat act as a barrier to gene flow between populations in these two areas as suitable boreal habitat does not exist in the Wyoming Basin. The authors also attributed strong phylogeographic structuring within the two clades to female philopatry and isolation by distance (Wilson et al. 2005). Additionally, red squirrels in the southern Rocky Mountains form a distinct clade from populations throughout the rest of the species' range (Arbogast et al. 2001). This divergent lineage most likely arose when populations in this region were isolated during times of coniferous forest fragmentation in the Pleistocene (Arbogast et al. 2001). In this species, both past and present geologic and climatic factors, as well as life history, have been implicated in structuring genetic patterns.

The American pika (*Ochotona princeps*) is a small mammal that lives in the Rocky Mountains of western North America. Like the red squirrel, the American pika is restricted to boreal habitat in alpine sky-islands or mountaintops. The American pika is also a poor disperser (Smith and Weston 1990). Due to these ecological traits, populations of this species have diverged into five lineages each endemic to a particular

mountain range (Galbreath et al. 2009). During glacial periods this species extended its range to lower elevations promoting admixture among now isolated populations within mountain ranges (Galbreath et al. 2010). The American pika and the red squirrel have specific habitat requirements that play a strong role in shaping their genetic structures.

The Mohave ground squirrel (*Xerospermophilus mohavensis*) has an extremely restricted geographic range in the northwestern corner of the Mojave Desert in the southwestern United States (Best 1995). This species shows much different phylogeographic structuring compared to its sister species, the round-tailed ground squirrel (*X. tereticaudus*). The round-tailed ground squirrel has a more extensive range than the Mohave ground squirrel, with its range encompassing the Mojave and Sonoran Deserts. The two species are sympatric in a narrow zone within the Mojave Desert. Bell et al. (2010) uncovered little phylogeographic structuring within the Mohave ground squirrel, but found four distinct lineages within the round-tailed ground squirrel. The authors attributed the lack of any phylogeographic structuring within the Mohave ground squirrel to high levels of ongoing gene flow and possibly a recent range expansion within the species' current distribution.

The pocket gopher (Genus *Thomomys*), a fossorial rodent in western North America, shows extreme genetic divergence among geographic regions, with many monophyletic groups having no distinct morphological differences among them (Belfiore et al. 2008, Ávarez-Castañeda 2010). Species of pocket gopher have low levels of dispersal and strong site-fidelity, ecological characteristics that allow for accumulation of genetic differentiation due to genetic drift within populations (Belfiore et al. 2008, Ávarez-Castañeda 2010).

Small mammals with restricted ranges or specific habitat requirements exhibit distinct genetic structuring across their ranges. However, small mammal habitat generalist species may show less genetic structuring than mammals bound to one type of habitat. The white-footed mouse exemplifies this within regions, though not across the entire species' range, which includes most of North America. In a phylogenetic study of the white-footed mouse in the mid-Atlantic region of the eastern United States, little gene flow and genetic structuring were detected among samples from New England southward to Tennessee and Georgia (Shipp-Pennock et al. 2005). But on a larger scale from New England west to Michigan and Minnesota, a clear genetic subdivision was present between eastern and western populations (Rowe et al. 2006). As within large mammals, small mammals can show different genetic patterns on different geographic scales. Phylogeographic patterns may be somewhat dependant on the size of the species' range. The white-footed mouse has a very large range and shows distinct phylogeographic structuring across large regions (Rowe et al. 2006), whereas the Mohave ground squirrel, which is restricted to a small region in the Mojave Desert, has very little phylogeographic structuring (Bell et al. 2010).

Small mammals in other regions show similar patterns to those in North America (Hewitt 2000, Faulkes et al. 2004, Grill et al. 2009, Krystufek et al. 2009a, Hürner et al. 2010). Most studies have focused on inferring how regional climatic and geologic history has shaped the genetic patterns within small mammal species, and few have focused on how unique life-history traits may have affected the phylogeography of small mammal species. Semi-aquatic small mammals have a more specialized ecology than other mammal species. The ability to move across water barriers may allow semi-aquatic

mammals to disperse farther than those that cannot, causing less genetic structuring across these species' distributions and creating phylogeographic patterns that are more similar to vagile large mammals. However, the habitat specialization of semi-aquatic rodents could restrict their dispersal patterns, and subsequently gene flow among populations, causing more distinct phylogeographic structuring than other small mammals.

Few phylogeographic studies of semi-aquatic small mammals have been conducted. Those species that have been studied show substantially different phylogeographic patterns. The Neotropical water rat (*Nectomys squamipes*) had significant genetic structuring detected by microsatellite markers (Almeida et al. 2005). This species rarely travels far from its stream habitat, causing restricted gene flow among different streams that would create genetic differentiation detectable even by quickly evolving microsatellite markers. On the other hand, the southern water vole (*Arvicola sapidus*), which inhabits the Iberian Peninsula and France, has a similar ecology to the Neotropical water rat, yet exhibited shallow mtDNA phylogeographic structuring across its range (Centeno-Cuadros et al. 2009).

The Eurasian beaver (*Castor fiber*), which is one of the largest rodents and has a more extensive range than the water vole, displayed major phylogeographic structuring across its range (Ducroz et al. 2005, Durka et al. 2005). The association of populations with specific rivers could have created this pattern with watersheds acting as barriers to gene flow. Alternatively, the pattern could be a consequence of the beaver's near extinction at the end of the 19th century when many populations were extirpated causing the species to undergo a drastic bottleneck. Genetic structuring has also been studied in

the North American beaver (*C. canadensis*; Crawford et al. 2009). However this study explored genetic patterns using microsatellites to survey recent gene flow between two populations in Illinois. The authors found the two populations to be genetically differentiated and suggested that gene flow was restricted between them by infrequent dispersal events.

The Eurasian otter (*Lutra lutra*) is a larger semi-aquatic mammal that has experienced similar anthropogenically caused population declines as the beaver. Contrary to the phylogeographic patterns found within the Eurasian beaver, the Eurasian otter displayed little to no geographic structuring based on both mtDNA and microsatellite markers possibly due to a recent origin of extant populations (Mucci et al. 2010). Few other semi-aquatic mammals have been studied in a phylogeographic context.

The Study Organism

The aquatic life history of the marsh rice rat (*Oryzomys palustris*) is crucial for interpreting the phylogeography of this species. The marsh rice rat is a medium-sized rodent that inhabits wetlands and salt marshes of the southeastern United States. It is one of the most common small mammals in healthy wetland ecosystems (Loxterman et al. 1998). The range of the marsh rice rat extends from southern New Jersey and southeastern Pennsylvania south through the Florida peninsula and west to southeastern Texas and northeastern Mexico. There also have been confirmed reports of the marsh rice rat north of its outlined range according to Wolfe (1982) in southern Kentucky, southern Illinois, and southwestern Missouri. The marsh rice rat appears to have

inhabited this general area since the Pleistocene epoch; the oldest fossils of the marsh rice rat, found in Florida and Georgia, date back to the early Sangamonian interglacial stage of the Pleistocene (120,000 bp; Wolfe 1982).

The marsh rice rat is adapted to a semi-aquatic lifestyle and unlike most terrestrial small mammals can easily disperse over water (Loxterman et al. 1998). Marsh rice rats have a water repellant pelage where air trapped between hairs helps them to stay afloat and also reduces heat loss to the water (Esher et al. 1978). They are good swimmers and can swim underwater for more than 10 meters (Hamilton 1946, Esher et al. 1978). They are omnivorous, eating aquatic organisms, insects, and wetland vegetation (Brunjes and Webster 2003, Kruchek 2004). Hamilton (1946) found that marsh rice rats were feeding mainly on salt marsh grass (*Spartina alterniflora*) and glasswort (*Salicornia europea*) in his Virginia study area, while sometimes eating snails and crustaceans. He also observed rice rats building feeding platforms from marsh grasses on which they fed during high tide when their habitat was flooded. Marsh rice rats will shift their diet seasonally concentrating on plant or animal food at different times of the year depending on the resource availability (Kruchek 2004). Though they have been seen active during the day, they are usually nocturnal (Hamilton 1946).

Systematics of the Genus Oryzomys

The genus *Oryzomys* is one of 28 genera assigned to the tribe Oryzomyini in the Cricetidae family of rodents (subfamily Sigmodontinae), one of the most speciose families of mammals (Musser and Carleton 2005). Sigmodontinae is an extremely diverse group with 84 genera and 377 species, making it an evolutionarily interesting

group for studies concerning speciation and diversity (Smith and Patton 1993, Musser and Carleton 2005, Weksler et al. 2006). Another unique characteristic of Sigmodontinae is its wide geographic distribution from the northeastern United States to southern Chile.

In the genus *Oryzomys* there are 43 described species, most of which inhabit South and Central America (Musser and Carleton 2005), though others name as many as 50 species (Nowak 1999). The genus *Oryzomys* is notoriously complicated and confusing, with many species originally placed in *Oryzomys* now assigned to separate genera (Musser and Carleton 1993, Bonvicino and Moreira 2001). Phylogenetic studies of *Oryzomys* have shown that the currently accepted morphologically based classification of species in this group does not accurately reflect their evolutionary relationships (Weksler 2006). Bonvicino and Moreira (2001) used cytochrome b to test the monophyly of three *Oryzomys* species groups in South America. These three groups were composed of morphologically similar species with overlapping ranges. The cytochrome b gene demonstrated that the genus *Oryzomys* is not monophyletic with respect to the genus *Nectomys*. Only one morphologically defined species group (*O. megacephalus*, *O. laticeps*, and *O. perenensis*) was found to be monophyletic. Similarly, Myers et al. (1995) found the genus *Oryzomys* to be paraphyletic using the cytochrome b gene. Weksler et al. (2006) described 10 new genera of Oryzomyine rodents and assigned only five species to the genus *Oryzomys* rendering the group monophyletic.

This unresolved and continual revision of the genus has weakened our understanding of the amount of diversity present in the *Oryzomys* group, as well as the evolutionary relationships among *Oryzomys* species. Recent molecular studies are

starting to uncover the nature of relationships within the Oryzomyini tribe (Musser and Carleton 1993, Bonvicino and Moreira 2001, Weksler et al. 2006).

These phylogenetic studies have focused exclusively on South and Central American *Oryzomys*. Three *Oryzomys* taxa inhabit North America and little attention has been given to uncovering the phylogenetic relationships within and among these species. Coues' rice rat (*O. couesi*) inhabits Central America, Mexico, and southeastern Texas. The marsh rice rat (*O. palustris*) is the northernmost *Oryzomys* species. As described above, the marsh rice rat inhabits the southeastern United States and is sympatric with the Coues' rice rat in southeastern Texas and northeastern Mexico. One other North American *Oryzomys* taxon, the silver rice rat (*O. argentatus*), is found only in the lower Florida Keys. The systematic status of this group is unresolved.

Research Objectives

My dissertation evaluated the genetic patterns within the marsh rice rat from the macroscale of phylogenetics, to the phylogeographic scale, to the microscale of population genetics. Chapter Two starts at the fundamental level of evaluating the systematic relationships of the six subspecies previously described within the marsh rice rat. Systematic designations within the marsh rice rat currently are based only on morphological data. This is problematic because environmental pressures and local habitat conditions can greatly influence the morphology of a species. This local adaptation and variability can obscure the evolutionary relationships among populations. Morphology is highly variably within populations of the marsh rice rat, sometimes more so than among populations (Goldman 1918). Females tend to group into two size classes

dependant on when they first breed; females who breed before they are fully mature are smaller in size (Paradiso 1960). The differences between these two groups exceed differences found between different geographic localities. Morphologically based taxonomic designations can be more subjective than ones made with the use of genetic markers. To distinguish among different evolutionary lineages, genetic data are imperative.

This study used mtDNA sequence data to infer the evolutionary relationships among the six marsh rice rat subspecies, as well as the relationship between the marsh rice rat and the silver rice rat. The systematic status of the silver rice rat in the Lower Florida Keys has been debated since the species description by Spitzer and Lazell in 1978 (Humphrey and Setzer 1989). Whether or not this morphologically and behaviorally unique group is a separate species remains unresolved. The marsh rice rat's sister species, Coues' rice rat (*O. couesi*), was recently confirmed as a separate species with DNA sequence data (Hanson et al. 2010). Coues' rice rat, which is found in Mexico and throughout Central America, had been classified as subspecific to the marsh rice rat, though more recently it had been treated as a separate species.

In Chapter Three, I thoroughly examined the phylogeographic patterns within the marsh rice rat. The southeastern United States harbors a vast amount of biodiversity, much of which is still being discovered (Odum 2002, Blaustein 2008, Graham et al. 2010). This area has been of biological interest for a long time because of its unique biotic assemblages and biogeographic history (Briggs et al. 1974). A diversity of habitats and extensive wetlands characterize the southeastern United States (Odum 2002). This geographic region is a climatic transitional zone, from temperate to subtropical. Many

temperate species' southern ranges end around Cape Canaveral, Florida on the east coast and Naples on Florida's west coast (Briggs et al. 1974).

Many species of the southeastern United States have been the subject of phylogeographic studies, including the earliest ones carried out by Avise et al (1979a and 1979b). Most of these studies have concluded that the Pleistocene glacial cycles shaped the present genetic structure of these species (Avise 1992, 1996). Present day biotic communities of the southeastern United States began to be formed towards the end of the most recent glacial period, the Wisconsin glaciation (18,000 bp; Delcourt 1993). At this time, the southeastern United States was mostly sand dune scrub and prairie along the coast all the way to the tip of the Florida peninsula (Gates 1993). To the north, temperate evergreen forests dominated the landscape. Eventually these forests spread south and became the major ecosystem that is seen in the southeast today (Gates 1993). After the Wisconsin glaciation when sea levels rose to their present position, wetlands formed along the eastern seaboard and Gulf coast (Gardner and Porter 2001). Many plants and animals are now dependent on this habitat type, including the marsh rice rat. Phylogeographic studies of species inhabiting the southeastern United States have hypothesized that the geologic and climatic history of the area has influenced different species similarly (Soltis et al. 2006). However, the uniqueness of the marsh rice rat, its habitat specialization and overwater dispersal ability, suggest that this species may show very different phylogeographic patterns.

Chapter Four considers the present day evolutionary processes that maintain the genetic structure and diversity of the marsh rice rat. Gene flow is a major evolutionary force; the marsh rice rat’s dispersal ability may allow for high levels of gene flow among

populations. Though populations may be structured geographically due to the distance among them in different regions, gene flow among local populations should influence levels of genetic diversity and differentiation.

This dissertation describes the genetic patterns within the marsh rice rat at all spatial scales. To make inferences about this species' evolutionary history, its entire genetic architecture should be examined, from the macroscale of phylogenetics, to a smaller scale of phylogeography, to the microscale of population genetics. The phylogenetic analyses of the marsh rice rat uncover the broader evolutionary relationships among subspecies. Studying the phylogeography of the marsh rice rat elucidates its geographic patterns of genetic variation. Genetic patterns among populations in different regions may shed light on how past climatic events affected this species. In this dissertation I identify the varying degrees of gene flow among populations, and quantify the levels of genetic diversity within populations. By identifying gene flow among populations, more recent patterns of genetic variation among populations in the same region can be described and the contemporary factors influencing genetic diversity within populations can be uncovered. From this genetic data, I propose conjectures about the evolutionary history of the marsh rice rat and place this species into the context of the biogeographic and geologic history of the southeastern United States.

Chapter Two

Systematics of the Marsh Rice Rat (*Oryzomys palustris*): The Presence of Distinct Evolutionary Lineages within a Morphologically Conservative Species

Background

The systematic relationship of the marsh rice rat, *Oryzomys palustris,* to other species in the Oryzomyini tribe, along with the taxonomic arrangement of the genus *Oryzomys,* has been controversial and problematic since the species was first described (Harlan 1837). The marsh rice rat's taxonomic relationships to other *Oryzomys* species have been revised many times (Baird 1857, Thomas 1893, Merriam 1901, Goldman 1918, Weksler 2003 and 2006, Weksler et al. 2006). As study of the marsh rice rat increased in geographic scope, unique populations were discovered and identified as subspecies. The naming of morphological subspecies has further complicated understanding of the marsh rice rat's evolutionary history.

The primary obstacle to identifying the evolutionary relationships in the genus *Oryzomys* is the morphological similarity found among taxa, which makes it difficult to identify separate species based on morphology alone (Bonvicino and Moreira 2001). As well as species of *Oryzomys* being morphologically similar, extensive morphological variation also can exist within a single species. Moreover, the amount of morphological variation present within a population of the marsh rice rat often can exceed the variation found between two populations from different geographic regions (Goldman 1918, Paradiso 1960). Genetic tools can now be used to uncover evolutionary relationships within *Oryzomys*, and may shed light on these taxonomic problems and controversies. Some genetic research has already been done to clarify and revise the systematic placement of currently and formerly recognized *Oryzomys* species in South and Central

America (Musser and Carleton 1993, Myers et al. 1995, Bonvicino and Moreira 2001, Weksler 2006), but little attention has been given to uncovering the systematics of the North American species using the genetic analyses that are now available.

Three *Oryzomys* taxa currently are recognized in North America: the marsh rice rat (*O. palustris*, Harlan 1837), Coues' rice rat (*O. couesi*, Alston 1876), and the silver rice rat (*O. argentatus*, Spitzer and Lazell 1978; Table 2.1). The marsh rice rat is the northernmost species of *Oryzomys*, inhabiting the eastern United States from southern New Jersey and southeastern Pennsylvania southward through the Florida peninsula and west along the Gulf Coast to southeastern Texas and extreme northeastern Mexico (Figure 2.1). This medium-sized rodent is semi-aquatic and lives in coastal and inland marsh habitat. The marsh rice rat is an excellent swimmer and has been recorded in a laboratory setting swimming underwater for 10 meters (Esher et al. 1978). They can also disperse across bodies of water; Forys and Moncrief (1994) observed individuals swimming distances of up to 300 meters.

Based on morphological analyses, six subspecies of the marsh rice rat are recognized: *O. p. coloratus* and *O. p. natator* in peninsular Florida; *O. p. planirostris* from Pine Island, Florida; *O. p. sanibeli* from Sanibel Island, Florida; *O. p. palustris* from New Jersey southward to northern Florida; and *O. p. texensis* in Louisiana and the Gulf Coast of Texas (Wolfe 1982; Figure 2.1, Table 2.1). All six differ in size and pelage color. For example, the Pine Island rice rat (*O. p. planirostris*) and the Sanibel Island rice rat (*O. p. sanibeli*) were both described by Hamilton (1955) as being smaller in size and having different colored pelages than *O. p. coloratus* and *O. p. natator* on the Florida mainland (Table 2.1).

A now extinct North American *Oryzomys* species inhabited the island of Jamaica until the late 1800's (Morgan 1993, Nowak 1999; Table 2.1). *Oryzomys antillarum* (Thomas 1898) was originally described as a separate species, and then later as a subspecies of the marsh rice rat, then as a subspecies of Coues' rice rat. This taxon is currently recognized as a separate species based on morphological divergence (Weksler et al. 2006). No other reports of *Oryzomys* inhabiting Caribbean islands exist, either in the fossil record or from present day sightings.

The evolutionary relationship of Coues' rice rat from Texas and Mexico to the marsh rice rat had been controversial until recently, with Coues' rice rat at some times recognized as a separate species (*O. couesi*) and other times classified as a subspecies of the marsh rice rat (*O. palustris aquaticus*; Harlan 1837, Baird 1857, Alston 1876, Thomas 1893, Meriam 1901, Goldman 1918, Hall 1960). They are morphologically divergent; Coues' rice rat is larger and less gray than the marsh rice rat. They also exhibit a different X chromosome structure (Benson and Gehlbach 1979, Wolfe 1982; Table 2.1). Schmidt and Engstrom (1994) studied sympatric populations of Coues' rice rat and the marsh rice rat in southeastern Texas and found no evidence of hybridization between the two groups based on allozyme analyses. Though collected at the same sampling sites, Coues' rice rat and marsh rice rat individuals were consistently different at many of the 31 allozyme loci studied. Individuals of each species were fixed for different alleles at four loci, and one or the other species had unique alleles at eight loci. Out of the 176 individuals collected at sites where both species lived sympatrically, none of them showed any genetic evidence for interbreeding. In a more recent study, DNA sequence data from the mitochondrial cytochrome b gene, a portion of exon 1 from the

nuclear interphotoreceptor retinoid binding protein gene, and intron 2 of the alcohol dehydrogenase 1 gene all supported the distinct species status of Coues' rice rat and the marsh rice rat (Hanson et al. 2010). Based on samples from throughout each species' range, these two species are reciprocally monophyletic.

The phylogenetic relationship of the silver rice rat (*O. argentatus*) from the lower Florida Keys to the mainland marsh rice rat (*O. palustris*) remains unresolved and controversial. The evolutionary relationship between these two groups is of special interest because the silver rice rat is considered highly endangered. Populations in the lower Florida Keys are protected by the U.S. Endangered Species Act, under which the silver rice rat is classified as a distinct vertebrate population. The silver rice rat is morphologically and behaviorally distinct. It has a silver-gray pelage and a more slender, narrower skull when compared to the mainland marsh rice rat (Spitzer and Lazell 1978, Goodyear 1991; Table 2.1). Silver rice rats have larger home ranges than mainland marsh rice rats and almost exclusively inhabit salt marsh (Spitzer 1983, Goodyear 1987, Goodyear 1992). They have a lower reproductive rate and are found at much lower densities than the mainland marsh rice rat (U.S. Fish and Wildlife Service 1999). Many studies over the last 30 years have argued whether or not this group is a separate species (Spitzer and Lazell 1978, Barbour and Humphrey 1982, Humphrey and Setzer 1989, Goodyear 1991). Morphological and genetic analyses have reached different conclusions. Based on morphology, Spitzer and Lazell (1978) first described the silver rice rat as a separate species (see also Goodyear 1991). Later Barbour and Humphrey (1982) argued that the silver rice rat could not have been separated from mainland populations long enough to undergo speciation and supported its subspecific relationship

to the mainland marsh rice rat. In a subsequent study, Humphrey and Setzer (1989) found the silver rice rat to be morphologically undifferentiated from the mainland marsh rice rat concluding that this taxonomic unit is neither a different species nor a distinguishable subspecies. However, a study comparing a small DNA region (291 base pairs) from the mitochondrial control region between individual rice rats from the Florida Keys (*O. argentatus*) and the Florida Everglades (*O. p. coloratus*) uncovered a unique control region haplotype, differing at 34 sites, within the silver rice rat, supporting the recognition of the lower Keys population as a distinct taxon (Gaines et al. 1997). However, the small amount of genetic differentiation between silver rice rats in the Florida Keys and marsh rice rats from the Florida Everglades, along with the silver rice rat's estimated short time of geographic isolation, led the authors to conclude that the silver rice rat is not a unique species. Recently, a nuclear DNA microsatellite genetic study concluded that the silver rice rat was adequately differentiated genetically to have subspecies status (Wang et al. 2005).

Interestingly, no rice rats currently inhabit the Upper Keys, separating the Lower Keys populations from populations on mainland Florida by more than 100 km (Goodyear 1987). Though the marsh rice rat has been observed swimming distances of up to 300 meters (Forys and Moncrief 1994), the geographic separation of the silver rice rat from the mainland marsh rice rat may be large enough to prevent gene flow between the two groups, thereby reproductively isolating the silver rice rat. Because the silver rice rat has been studied only since 1978, how long populations have not existed in the Upper Keys is unknown. The recent extirpation of the Upper Keys populations coinciding with urban development cannot be completely discounted. However, there is no scientific,

observational, or historical evidence of contact between the lower Keys' silver rice rat and mainland populations of the marsh rice rat. Based on sea level records, these populations could have been separated for a maximum of 3000 years (Barbour and Humphrey 1982). This amount of time may not have allowed for complete speciation. When sea level was lower than present day, during the last glacial maximum, populations in the Lower Keys were probably connected to populations in the Everglades via intermediate populations in the Upper Keys. When sea level rose with the melting of the glaciers, populations in the Upper Keys may have been extirpated, thereby isolating the Lower Keys populations from the mainland. Since this geographic isolation, the silver rice rat has been diverging, morphologically and genetically, from the mainland marsh rice rat.

The goal of this study was to clarify the systematic relationship between the marsh rice rat and silver rice rat, as well as use genetic data to test the validity of the morphological subspecies designations within the marsh rice rat. For this study I defined a species as an evolutionary significant unit (ESU) and a subspecies as a management unit (MU) as in Moritz 1994 and 2002. Accordingly, a species is a historically separated group of populations that is reciprocally monophyletic for mtDNA alleles. Moritz (1994) defines a subspecies or MU as a population with significant divergence of allele frequencies at nuclear or mitochondrial loci regardless of the phylogenetic distinctiveness of the alleles. Therefore, a subspecies may be within a paraphyletic clade due to incomplete lineage sorting (Avise 2004). If two groups have not been separated for an evolutionarily significant amount of time, yet there is little gene flow between them, ancestral polymorphisms can persist in both groups. In this scenario one group that may

be subspecific, appears to be in the same monophyletic clade as the other group, even though there is allelic divergence between the two indicating little to no gene flow. Over time the ancestral polymorphisms within the subspecies will disappear due to stochastic lineage sorting rendering that subspecific group monophyletic. Phylogenetic evidence, as well as geographic isolation, should be considered when naming subspecies (Lidicker 1962). Using many lines of evidence to delineate a species or subspecies strengthens these systematic, and somewhat arbitrary, assignments.

The current subspecies designations of the marsh rice rat are based on morphology; therefore using genetic data to investigate subspecies can provide further evidence and support for the recognition of previously defined morphological subspecies. Based on the mitochondrial cytochrome b gene, species differentiation in the closely related genus *Melanomys* ranges from 4.5% -- 7.6%, whereas individuals of the same subspecies are less than 2% divergent (Hanson and Bradley 2008). Cytochrome b divergence among other rodent species ranges from 1.3% -- 13% (average 7.3%; Baker and Bradley 2006). Intraspecific (subspecies) differentiation in other rodent groups ranges from 0 -- 4.7% in cytochrome b, with an average differentiation of 1.5% (Baker and Bradley 2006). Cytochrome b differentiation among populations within species ranges from 0 -- 1.4%, with an average differentiation of 0.6%. These aforementioned levels of genetic differentiation found within other rodents will be used in this study as a guide for identifying species, subspecies, and populations in the marsh rice rat, though exact values of differentiation are not unanimously agreed upon in the scientific community.

I hypothesized that because morphological differences do exist among the subspecies of the marsh rice rat, there will be genetic differences among them. I predicted that the six marsh rice rat subspecies each represent a distinct, monophyletic evolutionary lineage. I also predicted that the silver rice rat has not been separated from mainland populations long enough for genetic mutations to accumulate, therefore this group is not different enough to be a separate species. Given the morphological distinctions among subspecies and that they inhabit different geographic areas, the subspecies will exhibit some degree of genetic differentiation. Subspecies in Florida may be more differentiated genetically and have more variation than those in other regions of the species' range because four of the six subspecies exist only there.

Methods

Sample Collection

I collected tissue samples from marsh rice rats throughout the species' range. Samples from 36 localities were included to incorporate all morphologically distinct subspecies populations as well as samples of the silver rice rat and Coues' rice rat (Figure 2.1, Appendix A). Between one and 20 individuals were sampled from each population. In this study a population is a group of individuals occupying the same sampling area and adjacent habitats, or in the case of samples obtained from museums, individuals from the same county.

I obtained tissue samples by trapping individuals in Sherman live traps and cutting approximately 0.5 cm of tail tip from each animal captured using a pair of scissors. Individuals were weighed and sexed. Tissue samples were stored in 1.5 ml

screw cap tubes filled with a 20% DMSO (6 M NaCl) solution. Sampling methods were approved by the University of Miami Animal Care and Use Committee and followed methods approved by the American Society of Mammalogists Animal Care and Use Committee (Gannon et al. 2007). Additional samples were loaned from museum collections (tail tip, liver, or toe bone; Appendix A).

I included a total of 257 individuals of the marsh rice rat. This included individuals from all six subspecies of the marsh rice rat: *O. p. palustris* (n = 60), *O. p. texensis* (n = 140), *O. p. coloratus* (n = 20), *O. p. natator* (n = 4), *O. p. planirostris* (n = 8), and *O. p. sanibeli* (n = 12), as well as from the silver rice rat (n = 13). I also included nine individuals of Coues' rice rat. All samples and their specific localities are listed in Appendix A.

DNA Extraction and Mitochondrial DNA Sequencing

I isolated genomic DNA from tail tips and liver using a standard ethanol precipitation procedure. A DNeasy® tissue kit (Qiagen Inc., Valencia, California) was used to extract genomic DNA from museum toe bones. The mitochondrial cytochrome b gene (Cytb) and control region (CR) were amplified using the polymerase chain reaction (PCR; Saiki et al. 1988). Cytb is commonly used for systematic studies of rodents at the species level. The CR is ideal for intraspecific studies because it evolves more quickly than Cytb, therefore shows more variation at the population level (Bellinvia 2004). PCR primers for Cytb were forward MVZ05 – CGAAGCTTGATATGAAAAACCATCGTTG (Smith and Patton 1993) and reverse CB40 – CCACTAYCAGCACCCAAAGC (Hanson and Bradley 2008) and for the CR forward Ory5' – TACCATGAYCTTGTAAGTC (this

study) and reverse 2340-5 – GCATTTTCAGTGCTTTGC (Mendez-Harclerode et al. 2005).

For both Cytb and CR, the total PCR reaction volume was 10 μl, with 1 μl 10x buffer (2.5 mM $MgCl_2$ added), 1 unit Taq DNA polymerase, 0.1 mM dNTPs, and 14 pmol of each primer. The thermal profile for Cytb was: initial denaturation at 95°C (2 min), 30 cycles with denaturation at 95°C (45 s), annealing at 54°C (1 min), extension at 72°C (1 min 30 s), and a final extension at 72°C (8 min) (Hanson et al. 2010). The thermal profile for CR was: initial denaturation at 93.5°C (1 min), 33 cycles with denaturation at 93.5°C (40 s), annealing at 49°C (40 s), extension at 72°C (2 min 40 s), and a final extension stage at 72°C (2 min) (Mendez-Harclerode et al. 2005). Amplified fragments were purified using ExoSAP-IT enzymes (USB corp, Cleveland, Ohio) before cycle sequencing.

PCR fragments were sequenced using ABI Prism Big Dye Terminator v3.1 ready reaction mix (Applied Biosystems, Foster City, California). The primers used for initial PCR amplification were used with internal primers MVZ04 - GCAGCCCCTCAGAATGATATTTGTCCTC and MVZ45 - ACJACHATAGCJACAGCATTCGTAGG (Smith and Patton 1993) for Cytb and 500F - TCTCTTAATCTACCATCCTCCGTG (Castro-Campillo et al. 1999) and 1115 - ATGACCCTGAAGAARGAACCAG (Mendez-Harclerode et al. 2005) for the CR. Cycle sequencing was carried out using the following thermal profile: initial denaturation at 95°C for 1 min, then 40 cycles of denaturing at 95°C for 1 min, annealing at 50°C for 20 sec, and extension at 60°C for 4 minutes. I purified sequencing reactions using sephadex columns (Millipore), then dried them for 45 minutes with a vacuum centrifuge

and resuspended the reactions in 10 – 12 μl of Hi-Di Formamide (Applied Biosystems). Sequences were run on an ABI 3130xl automated sequencer (Applied Biosystems).

Nucleotide sequence chromatograms were aligned, edited, and proofed using Sequencher 4.6 software (GeneCodes, Ann Arbor, Michigan). Sequences for all individuals were then aligned in MEGA4 (Tamura et al. 2007). Aligned sequence files were imported into DNASP v.5 (Librado and Rozas 2009) to determine unique haplotypes.

MtDNA Sequence Data Analyses

I used these molecular data to investigate the phylogenetic relationship of the marsh rice rat to the silver rice rat and to examine the subspecific designations within the marsh rice rat. I calculated phylogenetic trees using only unique haplotypes found within each population. If the same haplotype was found in more than one population, that haplotypes was included for each population. Genetic distances were estimated among individuals from all subspecies and species and an analysis of molecular variance (AMOVA) was performed to calculate the molecular variance that is attributable to the subspecies designations (Excoffier et al. 1992).

Cytb and CR sequence data were used to estimate phylogenetic relationships among individuals. Because the Cytb gene and CR are on the mitochondrial genome, which is inherited as one unit, both were analyzed together for all phylogenetic analyses. Gaps in the CR alignment were coded using FASTGAP (Borchsenius 2007) using the conservative "simple indel coding" method described by Simmons and Ochoterena (2000). The nine Coues' rice rat individuals from Texas, Mexico, and Honduras were

included in phylogenetic analyses. The Chaco marsh rat (*Holochilius chacarius*) from Paraguay was used as the outgroup taxon in all analyses (Genbank accession numbers DQ227455 and AY863421).

I implemented maximum parsimony, maximum likelihood, and Bayesian analyses to estimate phylogenetic trees. Each analysis was performed at least twice to ensure the validity of the resulting trees. Parsimony analysis was conducted in PAUP v. 4.0b10 (Swofford 2002). Nucleotide positions were treated as equally weighted, unordered, discrete characters with four possible states: A, C, G, or T. The heuristic search method with tree bisection-reconnection branch swapping and 100 random addition replicates were used to estimate optimal trees. Nodal support of topologies was calculated using heuristic bootstrapping (BS) with 100 iterations (Felsenstein 1985). Searches were limited to 10,000 trees.

For maximum likelihood and Bayesian analyses the best-fit model of evolution was estimated for each mitochondrial region separately using the program MRMODELTEST (Nylander 2004). The most appropriate model of evolution for both Cytb and CR was the General Time Reversible model with parameters for invariant sites and rate variation (GTR + I + G) (Tavaré 1986). Maximum likelihood analysis was performed with the software program RAxML (Stamatakis 2006). CR and Cytb data were partitioned into separate regions (noncoding versus coding, respectively). Maximum likelihood support values were calculated with 100 bootstrap iterations (BS) using the rapid bootstrapping algorithm (Stamatakis et al. in preparation). A different random starting seed number was used for each run of the maximum likelihood analysis.

Bayesian analysis was carried out with the software program MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003). I implemented the site-specific gamma distribution and allowed for invariant sites. Cytb and CR regions were partitioned separately, with the Cytb coding region further partitioned by codon. I used four Markov-chains, 10 million generations, and a sample frequency of every 1,000th generation. The first 1,000 trees were discarded as "burnin" and a majority rule consensus tree was created with the remaining trees. Nodal support was calculated for tree topologies using clade posterior probabilities (PP) estimated with MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003).

I estimated the average genetic distances among all three North American *Oryzomys* species as well as among all six subspecies to assess taxonomic classifications. Because the four *O. p. palustris* samples from Mississippi and Tennessee clustered with *O. p. texensis* in all phylogenetic analyses, they were grouped with the *texensis* subspecies for this analysis. Genetic distances were estimated under the Kimura 2-parameter model of evolution (Kimura 1980) using MEGA4 (Tamura et al. 2007) and levels of genetic differentiation among all included taxonomic units were inferred. Genetic distances within each taxonomic unit were also estimated. Except as noted above, groups for comparison were determined a priori based on the morphological subspecies delimitations.

An analysis of molecular variance (AMOVA) was performed with the program ARLEQUIN to quantify genetic variation at three hierarchical levels: within populations, among populations within subspecies, and among subspecies (Excoffier et al. 1992, 2005). If the subspecies designations correctly portray the genetic structure of the marsh rice rat, I would expect most of the molecular variation to be at the subspecies level. This

would lend further support to the subspecies classifications being real biological and separate genetic entities within the marsh rice rat.

Results

I sequenced a total of 1143 base pairs for the cytochrome b gene (Cytb) and 1044 base pairs for the control region (CR). After combining the two data sets and aligning all sequences in MEGA4 (Tamura et al. 2000), there were a total of 2249 positions and 133 unique haplotypes. Gap coding with FASTGAP increased the number of informative characters by 52. Nucleotide frequencies for the Cytb sequences were A = 32.8%, C = 27.1%, G = 11.7%, and T = 28.4%, and for the CR A = 35.4%, C = 25%, G = 10.3%, and T = 29.3%. Transitions were 5.1 times more common than transversions in Cytb, and 2.36 times more common in the CR.

For the parsimony analysis, 495 informative characters were used to construct the limit of 10,000 equally most-parsimonious trees (length = 1375 steps, consistency index = 0.5505, retention index = 0.9567). A strict consensus analysis of all equally parsimonious trees revealed two major clades within the marsh rice rat (Clades A and B; Figure 2.2). Clade A included individuals from eastern populations including the subspecies *O. p. coloratus*, *O. p natator*, *O. p. sanibeli*, *O. p. planirostris*, and *O. argentatus* from Florida, as well as *O. p. palustris* from Delaware, Virginia, North Carolina, South Carolina, Georgia, and Alabama (Figure 2.3). Clade B included all individuals from the western subspecies *O. p. texensis*, as well as *O. p. palustris* individuals from Mississippi and Tennessee (Figure 2.4). Both clades were supported with the highest bootstrap support (100%). Many haplotypes within each clade formed

polytomies indicating that relationships could not be resolved possibly due to little sequence divergence among haplotypes.

The maximum likelihood and Bayesian analyses estimated the same major tree topology as the parsimony analysis, with Clades A and B being strongly supported (Figure 2.2; Clade A BS = 100%, PP = 1.0; Clade B BS = 96%, PP = 1.0). However, in the maximum likelihood phylogenetic tree, the node between Clades A and B was less strongly supported (BS = 89%) than it was in the parsimony and Bayesian analyses (BS = 100%, PP = 1.0). Within each clade, the maximum likelihood and Bayesian tree topologies were very similar to that resolved by parsimony analysis (Figures 2.5, 2.6). Polytomies were produced in the maximum likelihood analysis, though fewer than those found in the parsimony analysis. All minor clades within Clade A and within Clade B were the same in both the maximum likelihood and Bayesian trees. However, slight differences between the maximum likelihood and Bayesian analyses within Clade B existed (maximum likelihood tree not shown). Relationships among some clades differed. For example, a haplotype from Oklahoma and a haplotype from Tennessee grouped together, forming a monophyletic clade. In the Bayesian analysis this clade fell out separately from all other clades, but in the maximum likelihood analysis these two haplotypes grouped with haplotypes from Louisiana and Texas. In general, all three analyses produced the same results, with the Bayesian analysis resolving relationships among haplotypes the best.

Within the eastern Clade A both the Sanibel Island rice rat, *O. p. sanibeli*, (two haplotypes from 12 individuals) and the silver rice rat (three haplotypes from 13 individuals) were monophyletic (Figures 2.3 and 2.5). These haplotypes were unique to

these populations. However, the Sanibel Island rice rat and silver rice rat were both part of larger clades that included individuals from other Floridian subspecies. These clades were paraphyletic with all haplotypes from the Sanibel Island rice rat and silver rice rat being nested within broader phylogenetic groups. In both the maximum likelihood and Bayesian analyses, the silver rice rat was nested within a clade containing Everglades *O. p. coloratus* haplotypes. This association was weakly supported in the parsimony analysis; the relationship between the silver rice rat and other haplotypes could not be resolved. In all three analyses the Sanibel Island rice rat fell within a clade containing a haplotype from the Everglades (Miami-Dade County) and a haplotype from central Florida (Okeechobee County). The monophyletic grouping of the Sanibel Island rice rat individuals and the silver rice rat individuals were supported with the highest bootstrap and posterior probability support (BS = 100, PP = 1.0). The three *O. p. planirostris* haplotypes from the Pine Island, Florida subspecies did not form a clade, falling among individuals from the Everglades, *O. p. coloratus*, and central Florida, *O. p. natator*. Further, neither of these subspecies formed monophyletic clades.

Within western Clade B, haplotypes from Cameron County Texas, Willacy County Texas, and Tamaulipas Mexico formed a clade in all three analyses (BS = 100, PP = 1.0). These haplotypes are from extreme southeastern Texas and extreme northeastern Mexico. This clade could be an unidentified subspecies, though some haplotypes from Cameron County Texas also fell within other clades (Figures 2.4, 2.6). All three phylogenetic analyses supported that the boundary of the *O. p. texensis* group is farther east than originally thought, with *O. p. palustris* individuals from Mississippi and Tennessee grouping with *O. p. texensis* (clade B), instead of *O. p. palustris*. Wolfe

(1982) placed the geographic boundary between *O. p. texensis* and *O. p. palustris* along the Mississippi River between Mississippi and Louisiana. These data show that a genetic subdivision between *O. p. texensis* and *O. p. palustris* is actually present between Mississippi and Alabama farther east than the Mississippi River.

Genetic distances estimated using the Kimura 2-parameter model of evolution also highly supported a genetic separation between eastern and western marsh rice rat populations (Table 2.2). *Oryzomys p. texensis* in the west and the other marsh rice rat subspecies in the east were separated by a mean distance of 6.05% in Cytb and by a mean distance of 9.45% in the CR. Mean genetic distances among the eastern subspecies (*O. p. palustris*, *O. p. natator*, *O. p. coloratus*, *O. p. planirostris*, and *O. p. sanibeli*) ranged from 0.4% -- 1.2% for Cytb and from 0.9% -- 1.5% for the CR.

Genetic distances between the silver rice rat and the eastern marsh rice rat subspecies ranged from 0.5% -- 1.1% in Cytb and from 0.9% -- 1.4% in the CR, which is less than expected between separate species. However, the genetic distances between Coues' rice rat and *O. p. texensis* were 11.6% in Cytb and 13.1% in the CR, and between Coues' rice rat and the eastern marsh rice rat subspecies Kimura 2-parameter genetic distances averaged to be 10.9% for Cytb and 13.6% for the CR. These levels of genetic divergence support that the marsh rice rat and Coues' rice rat are separate species. Interestingly, *O. p. texensis*, which is sympatric with Coues' rice rat in southeastern Texas and northeastern Mexico, is genetically more differentiated from Coues' rice rat than it is from eastern marsh rice rat subspecies for Cytb, but less genetically differentiated from Coues' rice rat than from eastern marsh rice rat subspecies in the CR.

Kimura 2-parameter genetic distances within marsh rice rat subspecies ranged from 0.1% in *O. p. sanibeli* to 0.8% in *O. p. texensis* in Cytb, and from 0% in *O. p. sanibeli* to 1.5% in *O. p. texensis* in the CR. Genetic distance within the silver rice rat population was 0% for both mitochondrial DNA regions (Table 2.3).

The AMOVA analysis revealed that 66.26% of the genetic variation within the marsh rice rat is within populations, 29.04% is found among populations within subspecies, and only 4.7% of the genetic variation is explained by subspecies differences (Table 2.4). This analysis infers that little genetic variation can be attributed to the current subspecies designations. More genetic variation was found within and among populations of the marsh rice rat, than among putative subspecies.

Discussion

This study presents data that support the existence of a third rice rat species in North America, *O. texensis,* and three marsh rice rat subspecies, *O. p. palustris*, *O. p. sanibeli*, and *O. p. argentatus*. This study also provides additional genetic data supporting the species level genetic distinctness between the marsh rice rat and Coues' rice rat, confirming the results of Hanson et al. 2010. My hypothesis that there are genetic differences among subspecies of the marsh rice rat because morphological differences exist among them is supported in two of the six described subspecies (*O. p. palustris* and *O. p. sanibeli*). My original prediction that each marsh rice rat subspecies is a distinct evolutionary lineage is not supported by these genetic data. However, my prediction that the silver rice rat has not been separated from mainland populations long enough for genetic mutations to arise is strongly supported.

Oryzomys p. palustris, *O. p. natator*, *O. p. coloratus*, and *O. p. planirostris* do not form separate clades in all phylogenetic analyses. Individuals of these four subspecies group among each other, showing that little genetic divergence exists among them. The phylogenetic trees indicate that these four morphological subspecies are not genetically different subspecies; their haplotypes are not notably divergent nor are they historically separated populations reciprocally monophyletic for mtDNA alleles (Moritz 2002). The genetic distances among these four subspecies, compared to those among other rodent subspecies, do not support subspecific differentiation. In other rodents, Cytb genetic differentiation among populations ranges from 0 -- 1.4%, with a 0.6% average level of divergence (Baker and Bradley 2006). All four of these subspecies fall into the range of separate populations, not within the range of separate subspecies. The absence of genetic differentiation among these four subspecies may be due to high levels of gene flow among populations. Alternatively, if these populations are isolated from one another, they have only very recently become separated. These two population-level evolutionary processes could produce the shallow nodes evident among members of these four putative subspecies within the phylogenetic trees.

The Pine Island rice rat individuals were trapped in a restoration area on Little Pine Island, which is between Pine Island and mainland Florida. This mangrove coastal area was once the site of a sewage treatment plant where the predominant small mammal species was the black rat, *Rattus rattus* (Annette Nielsen, personal communication). Before restoration of this wetland habitat began in the mid-1990's, the land was dominated by invasive vegetation such as the Australian paper bark tree (*Melaleuca quinquenervia*), Australian pine (*Casuarina equisetifolia*), and the Brazilian pepper tree

(*Schinus terebinthifolius*), making the area uninhabitable to native wetland plant and animal species (Mariner Properties Development, Inc. Fort Myers, Florida). This area could have been recolonized by mainland marsh rice rats rather than by *O. p. planirostris* individuals from Pine Island after the wetland habitat became healthy enough to support native species. In future studies, specimens from Pine Island proper may still prove to be a distinct genetic subspecies. However, in the initial description of the Pine Island rice rat, Hamilton (1955) noted that individuals collected in North Ft. Myers Florida on the mainland did not differ from those he trapped on Pine Island, but that specimens from both locations differed from *O. p. natator* and *O. p. coloratus*. The fact that the island and mainland specimens were not differentiable from each other supports this study's finding that the Pine Island rice rat is not subspecific to the marsh rice rat. The initial description of the Pine Island rice rat as a separate subspecies was unwarranted.

The Sanibel Island rice rat, contrary to the other three eastern subspecies, did form a strongly supported clade. This group exhibits greater genetic divergence than the other eastern subspecies, and even slightly greater divergence than the silver rice rat. Intraspecific differentiation in other rodent groups ranges from 0 -- 4.7% divergence in Cytb, with an average of 1.5% (Baker and Bradley 2006). The Sanibel Island rice rat falls within this range and is slightly less than the 1.5% average. Sanibel Island is farther from mainland Florida than Pine Island (about 5 km versus 2 km), though Pine Island is just about 3 km north of Sanibel Island. Therefore, the Sanibel Island rice rat is slightly more geographically separated from other rice rat populations than the Pine Island rice rat. The amount of actual isolation between the Sanibel Island rice rat and other rice rats

warrants further study, especially because rice rats have been trapped on other small islands around Pine Island and Sanibel Island (Emily Woods, personal communication).

Based on all evidence currently documented for the silver rice rat, including its ecology, behavior, geographic isolation, and genetic differentiation, this population in the lower Florida Keys should be classified as the subspecies *O. p. argentatus*. Because the silver rice rat has only been separated from mainland populations for a maximum of 3000 years, this taxon has not significantly diverged enough from the marsh rice rat to justify its status as a separate species. At this point in time, the silver rice rat may be on a separate evolutionary trajectory, but if sea level was to fall, gene flow may be reestablished between the silver rice rat and marsh rice rat. Conversely if sea level rises, the silver rice rat population could go extinct due to habitat loss. The future relationship between the silver rice rat and the mainland rice rat is uncertain and highly dependent on geologic and climatic factors. Therefore, taxonomic decisions regarding this group should be conservative.

The silver rice rat does form a highly supported monophyletic group with unique haplotypes confirming some degree of taxonomic distinctness. However, based on genetic divergence, this taxonomic unit is only subspecific to the mainland marsh rice rat. Compared to the genetic differentiation between Coues' rice rat and the marsh rice rat (11.6% in Cytb, 13.6% in the CR), and even between eastern marsh rice rat populations and western marsh rice rat populations (6.05% in Cytb, 9.45% in the CR), there is insufficient genetic divergence to support the genetic species classification of the silver rice rat (separated from the marsh rice rat by 0.5% -- 1.1% in Cytb, 0.9% -- 1.4% in the CR). The silver rice rat does fall within the range for subspecific differentiation

compared to Cytb divergence in other rodents which ranges from 0 -- 4.7% in Cytb, with an average differentiation of 1.5% (Baker and Bradley 2006).

The taxonomic status of the silver rice rat could be argued either as a species or as a subspecies depending on the species definition used. Morphological and behavioral differences may be on par with species level distinctions. The environment of the Keys has affected certain traits of the silver rice rat, so that it is better adapted to living in mangrove habitats. Under the morphological species concept, the silver rice rat may be a separate species due to distinct morphological differences between it and mainland marsh rice rats (Spitzer and Lazell 1978, Goodyear 1991). The biological species concept states that a species is a group of interbreeding or potentially interbreeding individuals (Mayr 1942). Because the silver rice rat is geographically separated from the marsh rice rat, there is natural reproductive isolation between the two. Experiments to discover whether or not the two groups could interbreed would need to be conducted in a laboratory, as these two taxonomic units probably do not come into contact in nature. The phylogenetic species concept, with its many variations, generally defines a species as the smallest group of monophyletic individuals different from other such groups by at least one unique, fixed characteristic or diagnosable trait (Cracraft 1983, Baum 1992). According to this definition the silver rice rat is a separate species.

The taxonomic relationship of the silver rice rat to the marsh rice rat is transient in nature and has changed over evolutionary time. The findings of this study support some level of distinctiveness between the marsh rice rat and silver rice rat, but because these two groups have been separated for only a short amount of evolutionary time, I am being conservative in interpreting the genetic data produced by this study. Perhaps the silver

rice rat is a case of speciation in progress further supporting its continued protection by the United States Endangered Species Act and calling for continual study of this group.

Contrary to the silver rice rat, the western marsh rice rat subspecies (*O. p. texensis*) is as genetically divergent as some other rodent species and should be considered as a separate species, *O. texensis*. This was not a predicted result; *O. texensis* may represent a cryptic species. Identification of cryptic species with the use of genetic tools is not unusual within rodents; many genetic studies have uncovered significant divergence among morphologically similar taxa (Peppers and Bradley 2000, Riddle et al. 2000, Geise 2001). Nuclear sequence data from the interphotoreceptor retinoid-binding protein and the alcohol dehydrogenase 1 genes also support species level distinction between eastern and western marsh rice rats (Hanson et al. 2010). Previous studies of the marsh rice rat also found differentiation between eastern and western populations. In Allen's (1894) initial description of the subspecies *O. p. texensis*, western populations were described as much larger than *O. p. palustris* and very different in coloration compared to the other marsh rice rat subspecies (Table 2.1). Humphrey and Setzer (1989), who conducted a morphological analysis of the marsh rice rat, also were able to distinguish *O. p. texensis* from *O. p. palustris* based on some cranial measurements. However, they did not recognize western populations as a separate subspecies. Not only have past morphological studies found distinctions between eastern and western marsh rice rats, but also initial genetic analyses among North American *Oryzomys*. Schmidt and Engstrom (1994) identified unique, fixed alleles at three allozyme loci when comparing populations from Texas to a population in Georgia. Further analyses examining the

morphological, genetic, and possibly ecological distinctions between *O. texensis* and *O. palustris* will help to corroborate the taxonomic status of *O. texensis*.

Conclusions

This study has produced further questions about the intraspecific systematic relationships of the marsh rice rat and opened opportunity for more detailed studies of speciation in rodents, the most speciose group of mammals. The genetic division between the eastern and western clades, which occurs in the area between Mississippi and Alabama, needs to be studied in more detail. More intensive genetic sampling in this geographic area will clarify or further support this study's findings, and may help to infer what created and maintains this genetic pattern. From this current study, the presence of gene flow, or conversely reproductive isolation, across this genetic boundary cannot be resolved. Field studies of dispersal habits among populations in Mississippi and Alabama, as well as among populations throughout the marsh rice rat's entire range, are needed to understand the dynamics of gene flow between eastern and western populations. The ecology and behavior of the marsh rice rat has not been fully studied. Interpretation of the systematic relationships among North American *Oryzomys* taxa will be made clearer by relating genetic patterns to ecology and behavior. The next chapter of this dissertation (Chapter Three) examines these genetic relationships in more detail and relates geographic patterns of genetic diversity to geography and geologic history.

This system is interesting from an evolutionary perspective as the eastern and western clades may be undergoing speciation. Other groups within the marsh rice rat may also be under the influence of ongoing evolutionary processes, such as the silver rice

rat as discussed above, and the distinct clade of marsh rice rats in the most southwestern portion of this species' range. More genetic analyses among individuals from extreme southeastern Texas and northeastern Mexico should be conducted to investigate whether or not individuals from this area are an as yet undiscovered subspecies. Future studies of the marsh rice rat's ecology, behavior, and evolution, which are of particular interest due to this species' specialized habitat requirement, will depend on knowledge of the intraspecific systematic relationships uncovered in this study. The marsh rice rat's dependency on wetland habitat may have greatly shaped the phylogenetic relationships among different geographic groups.

Table 2.1. *Oryzomys* species and subspecies of North America. All measurements are reported in mm and weights in grams.

Species/Subspecies Name	Author of Initial Description	Type Locality	Geographic Range	Distinguishing Characteristics
Marsh rice rat (*O. palustris*)	Harlan 1837	Salem, Salem Co., New Jersey	Southeastern Pennsylvania and southern New Jersey to southern Florida and eastern Texas	Pelage, gray to grayish brown; length 226 – 305, tail 108 – 156, hind foot 28 – 37[1], skull 28.8[2]
O. p. coloratus	Bangs 1898	Cape Sable, Monroe Co., Florida	South Florida	Pelage, reddish[3]; length 283, tail 143.5, hind foot 33.4, skull 32.7[4]
O. p. natator	Chapman 1893	Gainesville, Alachua Co., Florida	Central Florida	Pelage, "darker"; larger than O. p. palustris length 286, tail 136, hind foot 33[5]
O. p. planirostris	Hamilton 1955	Pine Island, Lee Co., Florida	Pine Island and Little Pine Island, off the southwest coast of Florida, and mainland area 2 miles north of Ft. Myers, Florida	Pelage, brownish gray, "browner" than O. p. palustris, "lack of tawny coloration, smaller, weaker skull"; "size small", length 240, tail 122, hind foot 32, skull 30, weight 50.5 [6]
O. p. sanibeli	Hamilton 1955	Sanibel Island, Lee Co., Florida	Sanibel Island, off the southwest coast of Florida	Pelage amber brown and argus brown; "size small", length 263, tail 125, hind foot 33, skull 31.9, weight 71 [6]

O. p. palustris	Harlan 1837	Salem, Salem Co., New Jersey	Southeastern Pennsylvania and southern New Jersey to northern Florida and western Mississippi; north through Tennessee and into Kentucky	Same as *O. palustris*
O. p. texensis	Allen 1894	Rockport, Aransas Co., Texas	West of the Mississippi river to eastern Texas, including Oklahoma and Arkansas	"Large, pallid form"; Pelage yellowish gray-brown; length 277, tail 140, hind foot 30.5 [7]
Coues' rice rat (*O. couesi*)	Alston 1876	Coban, Guatemala	Southeastern Texas, throughout Mexico, into Central America	Larger, different x chromosome structure; Pelage dark brown; length 266.2, tail 135.8, hind foot 33.3, skull, weight 69.3 [8]
Silver rice rat (*O. argentatus*)	Spitzer and Lazell 1978	Cudjoe Key, Monroe Co., Florida	Lower Florida Keys, west of the Seven Mile Bridge	Pelage silver-gray; length 251 (♀), 259(♂), tail 121, 132, hind foot, 32, skull 29.3 30.4; narrower skull, more slender nasal bones[3]
Jamaican rice rat (*O. antillarum*) extinct	Thomas 1898	Jamaica	Jamaica	Pelage dorsum rufous, venter yellowish; length 250 mm, hind foot 29.2, skull 30.5 mm[9]

1 = Hall and Kelson 1959, 2 = Lowery 1974, 3 = Spitzer and Lazell 1978, 4 = Hamilton 1955, 5 = Chapman 1893, 6 = Hamilton 1955, 7 = Allen 1894, 8 = Benson and Gehlbach 1979, 9 = Allen 1942

Table 2.2. Average Kimura 2-parameter (Kimura 1980) genetic distances for two mitochondrial regions (cytochrome b gene and the control region) reported as percent divergence with standard error (computed by 1000 bootstrap replicates) between all taxonomic units.

Taxa	Cytochrome b	Control Region
O. p. palustris – *O. p. texensis*	5.9% ± 0.7%	9.3% ± 1.0%
O. p. coloratus	0.7% ± 0.1%	1.4% ± 0.2%
O. p. natator	0.6% ± 0.1%	1.3% ± 0.2%
O. p. planirostris	0.7% ± 0.1%	1.5% ± 0.3%
O. p. sanibeli	1.2% ± 0.3%	1.3% ± 0.3%
O. p. texensis – *O. p. coloratus*	6.0% ± 0.7%	9.4% ± 1.0%
O. p. natator	6.0% ± 0.7%	9.5% ± 1.0%
O. p. planirostris	6.1% ± 0.7%	9.6% ± 1.0%
O. p. sanibeli	6.1% ± 0.7%	9.3% ± 1.0%
O. p. sanibeli – *O. p. coloratus*	1.0% ± 0.3%	0.9% ± 0.2%
O. p. natator	0.9% ± 0.2%	0.9% ± 0.2%
O. p. planirostris	1.0% ± 0.3%	0.9% ± 0.3%
O. p. planirostris – *O. p. coloratus*	0.5% ± 0.1%	1.1% ± 0.2%
O. p. natator	0.5% ± 0.1%	1.1% ± 0.2%
O. p. coloratus – *O. p. natator*	0.4% ± 0.1%	1.1% ± 0.2%
O. couesi – *O. p. palustris*	11.0% ± 0.9%	13.5% ± 1.1%
O. p. texensis	11.6% ± 0.1%	13.1% ± 1.0%
O. p. coloratus	11.0% ± 0.9%	13.6% ± 1.1%
O. p. natator	10.9% ± 0.9%	13.5% ± 1.1%
O. p. planirostris	10.8% ± 0.9%	13.4% ± 1.1%
O. p. sanibeli	10.8% ± 0.9%	13.8% ± 1.2%
O. argentatus – *O. couesi*	10.7% ± 0.9%	13.6% ± 1.2%
O. p. texensis	5.9% ± 0.7%	9.6% ± 1.0%
O. p. palustris	0.7% ± 0.2%	1.4% ± 0.3%
O. p. coloratus	0.5% ± 0.2%	1.0% ± 0.2%
O. p. natator	0.5% ± 0.2%	1.1% ± 0.2%
O. p. planirostris	0.5% ± 0.2%	1.1% ± 0.3%
O. p. sanibeli	1.1% ± 0.3%	0.9% ± 0.3%

Table 2.3. Average Kimura 2-parameter (Kimura 1980) genetic distances for two mitochondrial regions (cytochrome b gene and the control region) reported as percent divergence with standard error (computed by 1000 bootstrap replicates) within all North American *Oryzomys* taxonomic units.

Taxa	**Cytochrome b**	**Control Region**
O. p. palustris	0.5% ± 0.1%	0.8% ± 0.2%
O. p. coloratus	0.5% ± 0.1%	0.1% ± 0.1%
O. p. texensis	0.8% ± 0.2%	1.5% ± 0.2%
O. p. natator	0.4% ± 0.1%	1.2% ± 0.3%
O. p. planirostris	0.4% ± 0.1%	0.7% ± 0.2%
O. p. sanibeli	0.1% ± 0%	0% ± 0%
O. couesi	2.8% ± 0.3%	3.8% ± 0.4%
O. argentatus	0%	0%

Table 2.4. Analysis of Molecular Variance (AMOVA) for the combined mitochondrial cytochrome b gene and control region sequence data from *Oryzomys palustris*.

Source of Variation	df	Sum of Squares	Variance Components	Percent Variation
Among subspecies	6	15.784	0.02372 (Va)	4.7
Among populations Within subspecies	25	34.008	0.14667 (Vb)	29.04
Within populations	225	75.286	0.33460 (Vc)	66.26
Total	256	125.078	0.50500	

Figure 2.1. Distribution of the marsh rice rat (*Oryzomys palustris*), Coues' rice rat (*Oryzomys couesi*) in North America. Distributions of marsh rice rat subspecies are delineated by dotted lines. Geographic localities of samples included in this study are marked with a closed circle (*Oryzomys palustris*) or an open square (*Oryzomys couesi*). Map modified from Hanson et al. 2010, Hall 1981, and Humphrey and Setzer 1989.

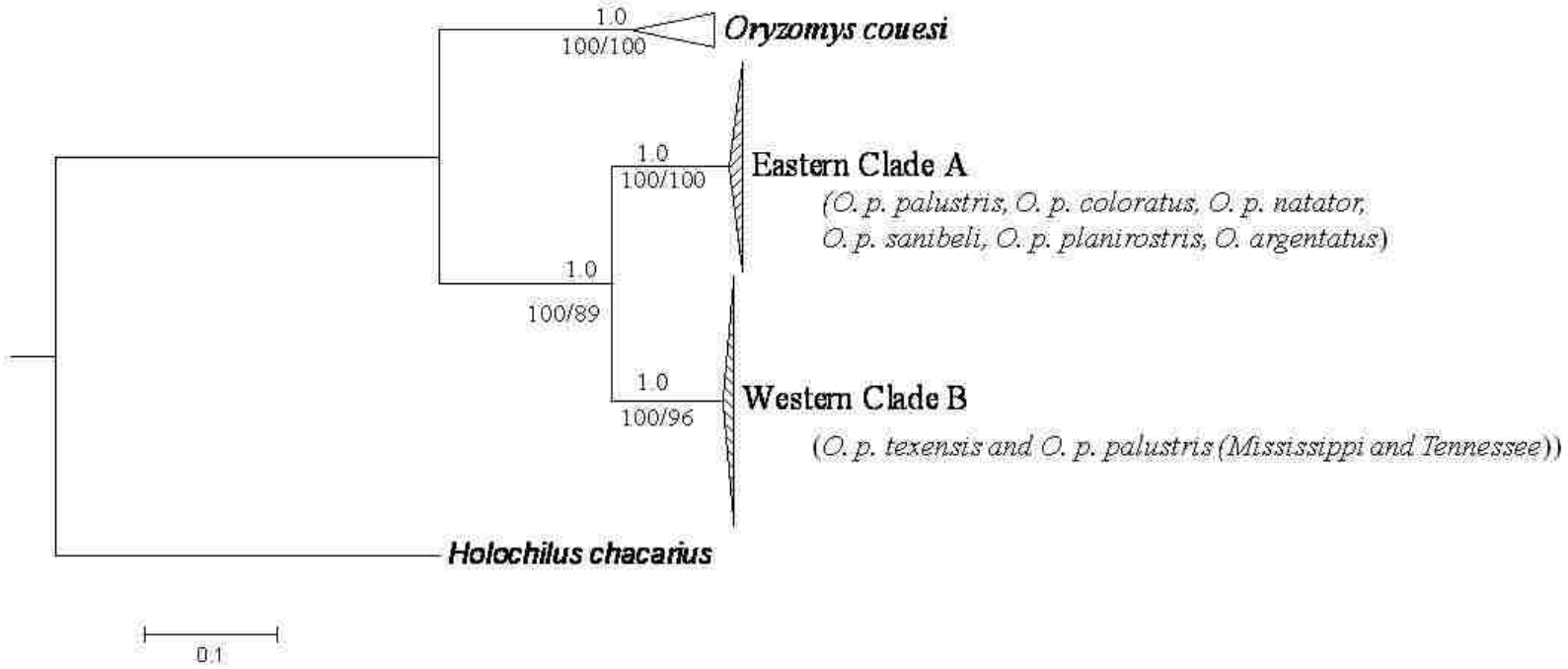


Figure 2.2. Phylogenetic tree estimated by Bayesian, maximum likelihood, and parsimony analyses of *Oryzomys palustris* mitochondrial Cytochrome b and control region sequence data. All three analyses calculated the same topology. Posterior probabilities greater than 0.95 are given above nodes and bootstrap values greater than 60% are given below nodes (parsimony bootstrap values/maximum likelihood bootstrap values). Major clades were collapsed to clarify major relationships among clades.

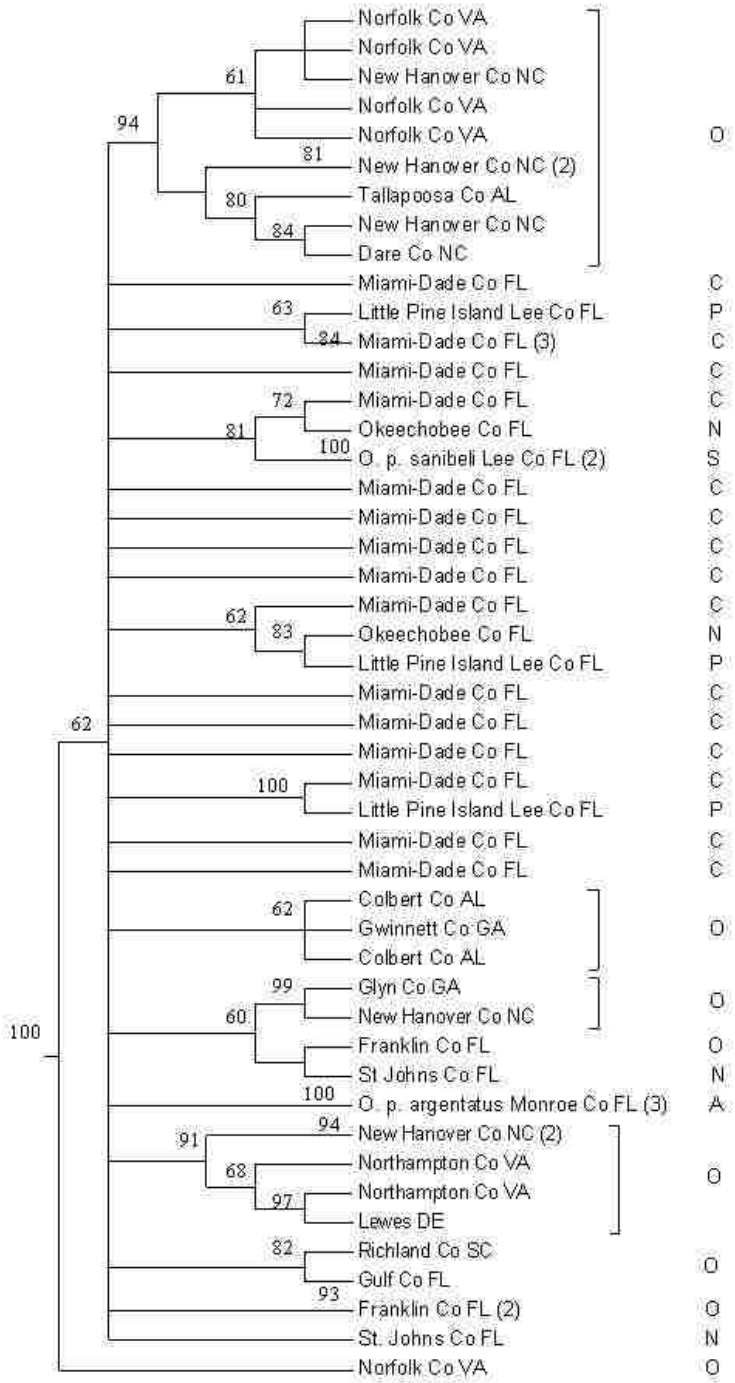


Figure 2.3. Strict consensus analysis of 10,000 most parsimonious trees – Eastern Clade A. Eastern Clade A includes individuals assignable to *Oryzomys palustris palustris* (O), *O. p. natator* (N), *O. p. coloratus* (C), *O. p. planirostris* (P), *O. p. sanibeli* (S), and *Oryzomys argentatus* (A). Only *O. p. sanibeli* (S) and *O. argentatus* (*O. p. argentatus*, A) form strongly supported monophyletic clades and should be classified as subspecies. Numbers in parentheses refer to the number of haplotypes within that collapsed clade. Bootstrap values greater than 60% are given. Nodes with bootstrap values greater than 80% are considered strongly supported, and values less that 70% are considered low.

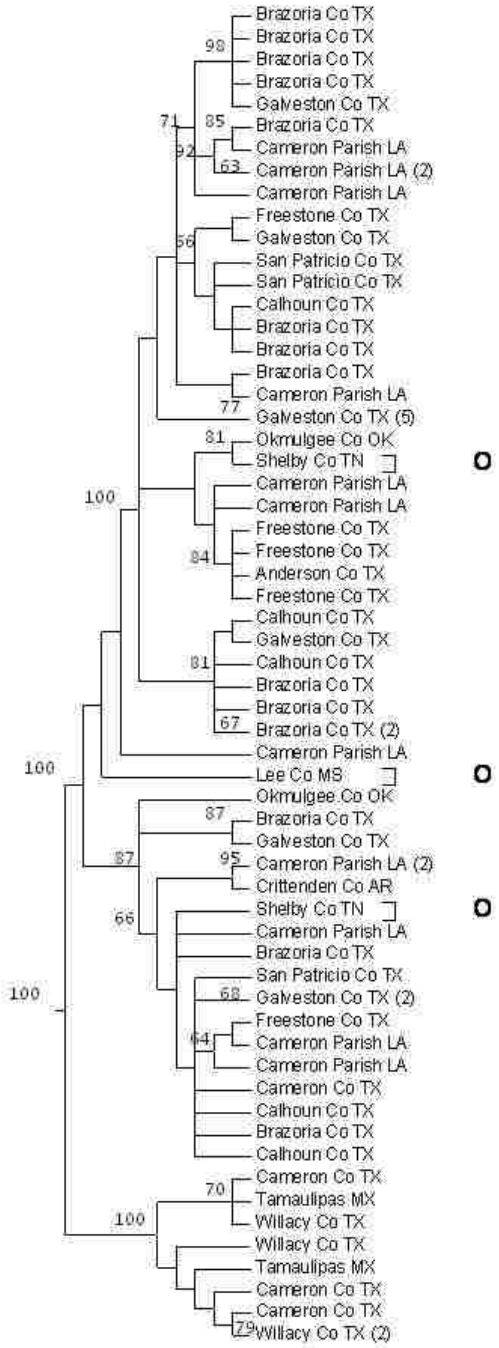


Figure 2.4. Strict consensus analysis of 10,000 most parsimonious trees – Western Clade B. This clade includes individuals assignable to *Oryzomys palustris texensis* and *O. p. palustris* individuals from Mississippi and Tennessee. Because of the amount of divergence between Clade A and Clade B, Clade B should be assigned the species name *Oryzomys texensis*. Haplotypes followed by an O are originally assignable to *O. p. palustris*. All other haplotypes are *O. p. texensis*. Numbers in parentheses refer to the number of haplotypes within that clade. Bootstrap values greater than 60% are given. Nodes with bootstrap values greater than 80% are considered strongly supported, and values less that 70% are considered low.

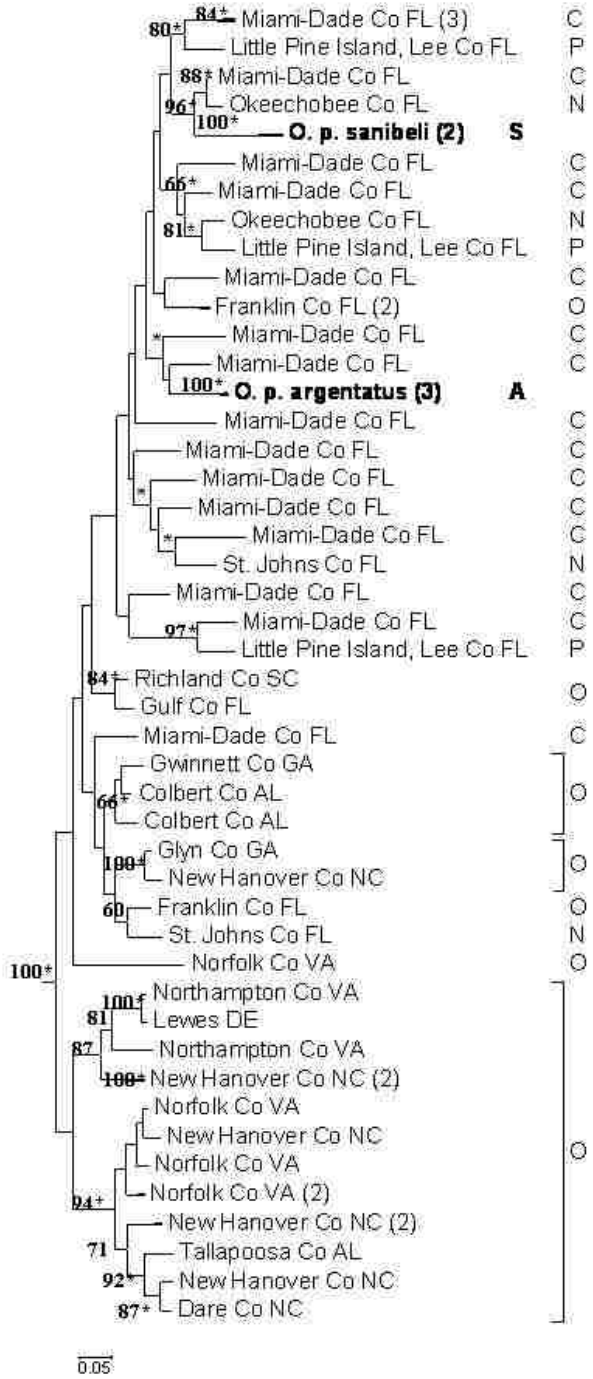


Figure 2.5. Maximum likelihood and Bayesian analyses – Eastern Clade A. This clade includes individuals assignable to *Oryzomys palustris palustris* (O), *O. p. natator* (N), *O. p. coloratus* (C), *O. p. planirostris* (P), *O. p. sanibeli* (S), and *Oryzomys argentatus* (A). Only *O. p. sanibeli* (S) and *O. p. argentatus* (A) form strongly supported monophyletic clades and should be classified as subspecies. Numbers in parentheses refer to the number of haplotypes within that collapsed clade. Bootstrap values greater than 60% are given at nodes. Posterior probabilities higher than 0.95 are indicated with an asterisk.

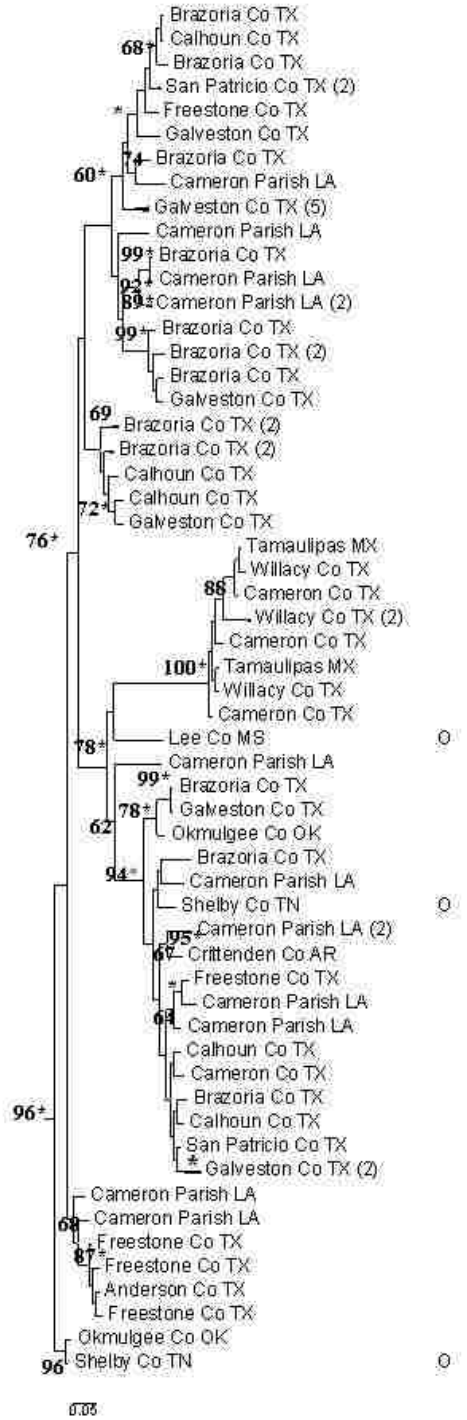


Figure 2.6. Maximum likelihood and Bayesian analyses – Western Clade B. This clade includes individuals assignable to *Oryzomys palutris texensis* and *O. p. palustris* individuals from Mississippi and Tennessee. Haplotypes followed by an O were originally assignable to *O. p. palustris*. All other haplotypes are *O. p. texensis*. Numbers in parentheses refer to the number of haplotypes within that collapsed clade. Bootstrap values greater than 60% are given at nodes. Posterior probabilities higher than 0.95 are indicated with an asterisk.

Chapter Three

Phylogeographic Patterns within the Marsh Rice Rat (*Oryzomys palustris*) Based on Mitochondrial DNA Markers

Background

Phylogeography seeks to relate intraspecific evolutionary relationships within species to the geography of species distributions (Avise 2000, Hickerson et al. 2010). The evolutionary processes that shaped a species present genetic structure can be inferred by correlating phylogeographic patterns with past geologic and climatic history. Phylogeography can detect hybrid zones, isolation, and major genetic discontinuities (Swenson and Howard 2005). Phylogeographic breaks often coincide with geographic barriers to gene flow, but many times a past geographic or climatic event better explains the presence of major genetic disjunctions within species. Often vicariant events, or allopatric speciation, are hypothesized to explain major genetic patterns. During the Pleistocene glaciations, populations may have genetically diverged in isolated glacial refugia (Hewitt 2000 and 2001). Thus phylogeography gives insight to the process of speciation and offers expainations for regional biogeographic patterns.

Phylogeographic studies of small mammals have shown many species to have spatially structured genetic patterns (Hayes and Harrison 1992, Riddle et al. 1993, Van Vuuren and Robinson 1997, Santucci et al. 1998, Arbogast et al. 2001, Wilson et al. 2005, Ávarez-Castañeda 2010, Bell et al. 2010, Galbreath et al. 2010). Deep genetic subdivisions can exist across a species' distribution (Avise 2000). This pattern is shaped by their small home ranges, mating systems, and dispersal patterns. In contrast, 1arger mammals have more variable phylogeographic patterns; some have little or no apparent

structuring, while others exhibit substantial genetic structuring (Cronin 1992, Morin et al. 1994, Ramey 1995, Clegg et al. 1998, Vìla et al. 1999, Comstock et al. 2002, Hundertmark et al. 2002, Zenger et al. 2003). Larger mammals are expected to show different patterns of genetic diversity than smaller species because they are generally more vagile and have larger home ranges. However, larger mammals may still have concordant genetic patterns with smaller mammals if past geologic and climatic factors have been a strong force in shaping the landscape in which these species exist. Phylogeographic patterns in mammals usually are concordant with behavior, natural history, and both past and present regional geology and climate.

The marsh rice rat (*Oryzomys palustris*) presents a unique opportunity for a phylogeographic study because of its close association with wetlands. This small mammal is adapted to a semi-aquatic lifestyle and can disperse between islands over open water (Forys and Moncrief 1994, Loxterman et al. 1998). This species' dependency on wetlands and its ability to disperse over water for longer distances than other sympatric small mammal species, such as the white-footed mouse (*Peromyscus leucopus*; Forys and Moncrief 1994), and the cotton rat (*Sigmodon hispidus*; Esher et al. 1978), could be significant factors in shaping the marsh rice rat's present genetic structure and phylogeographic patterns. The marsh rice rat's swimming ability may have created phylogeographic patterns that appear more like those in larger mammals. The marsh rice rat could show less genetic structuring than other small mammals due to its higher vagility. However, this species' restriction to wetland habitat could cause more phylogeographic structuring than expected for a vagile mammal, especially if wetland

habitat is highly fragmented. Regionally, marsh rice rat populations should show little genetic divergence, but among regions there may be phylogeographic structuring.

The southeastern United States has been of biological interest for a long time because of its unique climate and biotic assemblages, as well as the strong influence of glacial sea level changes upon the region's biogeography (Briggs et al. 1974). The topography of the Gulf and Atlantic coastal areas has changed considerably with variations in sea level throughout the Pliocene and Pleistocene (Leigh 2008, Morgan and Emslie 2010). Sea level is estimated to have fluctuated from 50 m above and 100 m below present day sea level (Emslie 1998). When sea level rose coastal areas were inundated, isolating upland habitats and forming small islands along the coast. The Florida peninsula became a series of islands and was isolated from the mainland during these times of higher sea level (Fairbridge 1974). During periods of lower sea level, previously disconnected coastal areas became continuous landscapes and the area of the Florida peninsula was much greater than present day. These sea level changes caused habitats to expand and contract, altering the landscape by forming barriers to dispersal in previously connected habitats or by creating corridors between historically separated areas (Morgan and Emslie 2010).

The North American ice sheet during the glacial periods of the Pleistocene also influenced the biogeography of the southeastern United States, though the glacier itself did not extend past present day New York and New Jersey (Dyke et al. 2002). Glaciers significantly cooled the climate, which forced species southward and altered floral assemblages (Gates 1993). Glacial maxima were characterized by drier conditions; the southeastern United States, and especially the Florida peninsula, is thought to have been

more arid and therefore less favorable for marshy habitats (Delcourt 1993). However, when the glaciers retreated and the climate warmed, coastal areas were characterized by extensive marsh habitat (Emslie 1998).

The first phylogeographic studies were conducted on species inhabiting the southeastern United States, and the phylogeographic patterns of many plants and animals inhabiting this region have been studied (Soltis et al. 2006). From these studies, authors have made inferences about the past biogeography of this region. Distinct phylogeographic structure has been observed in the southeastern United States for many unrelated vertebrate taxa. Within most species, eastern and western populations are genetically differentiated (Avise et al. 1979a, Bermingham and Avise 1986, Avise and Nelson 1989, Vogler and DeSalle 1993, Phillips 1994, Osentoski and Lamb 1995, Walker et al. 1995, Donovan et al. 2000, Lemmon et al. 2007, Pauly et al. 2007, Douglas et al. 2009, Jackson and Austin 2009, Fontanella and Siddall 2010). In many of these species, the split between lineages seems to consistently occur around the Apalachicola River area of northern Florida reflecting a historical barrier to gene flow (Avise 1992). This genetic division may be a pattern that arose during the Pleistocene glacial maxima, when populations of many species inhabited separate eastern and western refugia. For example, the eastern woodrat (*Neotoma floridana*) exhibits distinct mitochondrial DNA (mtDNA) clades in the north, south, and west with different levels of genetic variation within each clade (Hayes and Harrison 1992). The authors attribute this pattern to vicariant events caused by Pleistocene glacial cycles. Similarly, the white-tailed deer (*Odocoileus virginianus*) is geographically structured into three distinct groups of mtDNA haplotypes (Ellsworth et al. 1994). The Apalachicola River separates these

species' eastern and western mtDNA clades. The genetic patterns of the white-tailed deer and the eastern woodrat are spatially concordant with the variation observed in other unrelated species inhabiting the southeastern United States such as freshwater fish (*Amia calva*, *Lepomis punctatus*, and *Lepomis microlophus*; Bermingham and Avise 1986), the tiger beetle (*Cicindela dorsalis*; Vogler and DeSalle 1993), gopher tortoise (*Gopherus polyphemus*; Osentoski and Lamb 1995), American alligator (*Alligator mississippiensis*; Davis et al. 2002), salamanders (*Ambystoma talpoideum*; Donovan et al. 2000 and *Pseudobranchus striatus*; Liu et al. 2006), and the musk turtle (*Sternotherus minor*; Walker et al. 1995). More than 20 animal and plant species have been documented to share a genetic discontinuity at the Apalachicola River (Soltis et al. 2006). This phylogeographic concordance supports the hypothesis that species in the southeastern United States have been influenced by the same historical climatic and geologic events.

Not all species studied in the southeastern United States exhibited a genetic discontinuity at the Apalachicola River. An Appalachian Mountain discontinuity was uncovered in salamander species of the southeastern United States (*Ambystoma tigrinum tigrinum*, Church et al. 2003, and *Ambystoma maculatum*, Donovan et al. 2000, Zamudio and Savage 2003). Like the Apalachicola River genetic discontinuity, the genetic disjunction at the Appalachian Mountains has been attributed to the presence of separate refugia on opposite sides of the Appalachians (Soltis et al. 2006). The Mississippi River also is an east-west genetic divide between species such as the North American bullfrog (*Rana catesbiana* Austin et al. 2004), the northern Leopard frog (*Rana pipiens*; Hoffman and Blouin 2004), and the northern short-tail shrew (*Blarina brevicauda*; Brant and Ortí 2003). This too is attributed to eastern and western populations having inhabited separate

refugia on either side of the Mississippi River. Other major rivers in the southeastern United States have been implicated as the site of major phylogeographic breaks, such as the Tombigbee River and Savannah River (Soltis et al. 2006, Degner et al. 2010).

In this study I explored the phylogeographic patterns within the marsh rice rat in the southeastern United States in order to infer how the climatic and geologic history of this region has influenced this species. I also compared the phylogeography of the marsh rice rat to other species of the southeastern United States. I hypothesized that the marsh rice rat was shaped by the same climatic and geologic events that affected other species in the region. I predicted that the present distribution of genetic diversity is geographically structured within the marsh rice rat. Populations in the northeastern region of their distribution (Delaware, Virginia, North Carolina, and South Carolina), in the southwestern region (Texas, Oklahoma, Arkansas, Louisiana, and Mexico), and in the southeastern region (Florida, Georgia, Mississippi, and Alabama) will each form a clade within the intraspecific phylogeny due to the influence of past climatic and geologic factors. Populations in these three geographic areas are descended from ancestral populations which may have been geographically isolated during Pleistocene glacial periods. As in most other species studied in the southeastern United States, mtDNA haplotypes of the marsh rice rat will also be divided into eastern and western groups. I predicted that this genetic division would be at the Apalachicola River which flows south along the border between Alabama and Georgia into the Florida Panhandle. However, the marsh rice rat may have a less distinct separation between eastern and western populations than other vertebrates because they are more vagile than these other species that show this east-west division.

Four of the six morphological subspecies occur only in Florida, which led me to hypothesize that Floridian marsh rice rats are more genetically differentiated from populations in the other regions of their distribution. Populations on the Florida peninsula may have been partially isolated from mainland populations during the sea level rise before the last glacial period. The unique dispersal ability of the marsh rice rat may have allowed for gene flow between the Floridian peninsular populations and mainland populations, albeit at a reduced level compared to when sea level was lower.

Methods

Sample Collection

I used the same sampling procedure as I did for the systematics study presented in Chapter Two. Tissue samples were collected from marsh rice rats throughout the species' range. Samples from 32 localities were included to incorporate all geographic regions of the marsh rice rat's range (Figure 3.1). Between one and 20 individuals were sampled from each population for a total of 257 individuals (Appendix A). In this study a population is a group of individuals occupying the same sampling area and adjacent habitats or in the case of samples obtained from museums, individuals from the same county.

Tissue samples were obtained by trapping individuals in Sherman live traps and cutting approximately 0.5 cm of tail tip from each animal captured using a pair of scissors. Tissue samples were stored in 1.5 ml screw cap tubes filled with a 20% DMSO (6 M NaCl) solution. Sampling methods were approved by the University of Miami Animal Care and Use Committee and followed methods approved by the American

Society of Mammalogists Animal Care and Use Committee (Gannon et al. 2007). Samples were also loaned from museum collections (tail tip, liver, or toe bone, Appendix A).

DNA Extraction and Mitochondrial DNA Sequencing

Genetic data collection for this phylogenetic study was the same as that I implemented in the systematics study described in Chapter Two. Genomic DNA was isolated from tail tips and liver using a standard ethanol precipitation procedure. A DNeasy® tissue kit (Qiagen Inc., Valencia, California) was used to extract genomic DNA from museum toe bones. The mitochondrial cytochrome b gene (Cytb) and control region (CR) were amplified using the polymerase chain reaction (PCR; Saiki et al. 1988). Cytb is commonly used for phylogenetic studies of rodents and the CR, the main non-coding region in the mitochondrial genome is ideal for phylogeographic investigations. The CR is highly variable and evolves ten times faster than the rest of the mitochondrial genome. Therefore, this mtDNA region shows more variation at the population level (Bellinvia 2004). PCR primers for Cytb were forward MVZ05 – CGAAGCTTGATATGAAAAACCATCGTTG (Smith and Patton 1993) and reverse CB40 – CCACTAYCAGCACCCAAAGC (Hanson and Bradley 2008), and for the CR forward Ory5' – TACCATGAYCTTGTAAGTC (this study) and reverse 2340-5 – GCATTTTCAGTGCTTTGC (Mendez-Harclerode et al. 2005).

For both Cytb and CR, the total PCR reaction volume was 10 μl, with 1 μl 10x buffer (2.5 mM $MgCl_2$ added), 1 unit Taq DNA polymerase, 0.1 Mm dNTPs, and 14 pmol of each primer. The thermal profile for Cytb was: initial denaturation at 95^{o}C (2

min), 30 cycles with denaturation at 95°C (45 s), annealing at 54°C (1 min), extension at 72°C (1 min 30 s), and a final extension at 72°C (8 min) (Hanson et al. 2010). The thermal profile for CR was: initial denaturation at 93.5°C (1 min), 33 cycles with denaturation at 93.5°C (40 s), annealing at 49°C (40 s), extension at 72°C (2 min 40 s), and a final extension stage at 72°C (2 min) (Mendez-Harclerode et al. 2005). Amplified fragments were purified using ExoSAP-IT enzymes (USB corp, Cleveland, Ohio) before cycle sequencing.

PCR fragments were sequenced using ABI Prism Big Dye Terminator v3.1 ready reaction mix (Applied Biosystems, Foster City, California). The primers used for initial PCR amplification were used with internal primers MVZ04 - GCAGCCCCTCAGAATGATATTTGTCCTC and MVZ45 - ACJACHATAGCJACAGCATTCGTAGG (Smith and Patton 1993) for Cytb and 500F - TCTCTTAATCTACCATCCTCCGTG (Castro-Campillo et al. 1999) and 1115 - ATGACCCTGAAGAARGAACCAG (Mendez-Harclerode et al. 2005) for CR. Cycle sequencing was carried out using the following thermal profile: initial denaturation at 95°C for 1 min, then 40 cycles of denaturing at 95°C for 1 min, annealing at 50°C for 20 sec, and extension at 60°C for 4 minutes. Sequencing reactions were purified using sephadex columns (Millipore), then dried for 45 minutes with a vacuum centrifuge and resuspended in 10 – 12 μl of Hi-Di Formamide (Applied Biosystems). Sequences were run on an ABI 3130xl automated sequencer (Applied Biosystems).

Nucleotide sequence chromatograms were aligned, edited, and proofed using SEQUENCHER 4.6 software (GeneCodes, Ann Arbor, Michigan). Sequences for all individuals were then aligned in MEGA4 (Tamura et al. 2007). Aligned sequence files

were imported into DNASP v.5 (Librado and Rozas 2009) to determine unique haplotypes.

Phylogenetic Analyses

I used this molecular data to investigate the phylogeography of *O. palustris*. Intraspecific phylogenetic trees were estimated using Cytb and CR sequence data to uncover the genetic relationships among populations from different geographic regions and within geographic regions. The same phylogenetic analyses that were carried out to examine the systematics of the marsh rice rat in Chapter Two were used again in this study to examine phylogeographic patterns within this species. Because the Cytb gene and CR are on the mitochondrial genome, which is inherited as one unit, both were analyzed together for all phylogenetic analyses. Gaps in the CR alignment were coded with FASTGAP (Borchsenius 2007) using the conservative "simple indel coding" method described by Simmons and Ochoterena (2000). *Oryzomys couesi* from Honduras, Mexico, and Texas (n = 9), and *Holochilius chacarius* from Paraguay (n = 1) were used as the outgroup taxa in these analyses (Genbank accession numbers for *Holochilius chacarius*: DQ227455 and AY863421, for *Oryzomys couesi* see Appendix A).

Maximum parsimony, maximum likelihood, and Bayesian analyses were used to estimate phylogenetic trees. Each analysis was performed at least twice to ensure the validity of the resulting trees. Parsimony analysis was conducted in PAUP v. 4.0b10 (Swofford 2002). Nucleotide positions were treated as equally weighted, unordered, discrete characters with four possible states: A, C, G, or T. The heuristic search method with tree bisection-reconnection branch swapping and 100 random addition replicates

were used to estimate optimal trees. Nodal support of topologies was calculated using heuristic bootstrapping (BS) with 100 iterations (Felsenstein 1985). Searches were limited to 10,000 trees.

For maximum likelihood and Bayesian analyses the best-fit model of evolution was estimated for each mitochondrial region separately using the program MRMODELTEST (Nylander 2004). The most appropriate model of evolution for both Cytb and CR was the General Time Reversible model with parameters for invariant sites and rate variation (GTR + I + G) (Tavaré 1986). Maximum likelihood analysis was performed with the software program RAxML (Stamatakis 2006). Cytb and CR data were partitioned into separate regions (coding versus noncoding respectively). Maximum likelihood support values were calculated with 100 bootstrap (BS) iterations using the rapid bootstrapping algorithm (Stamatakis et al. in preparation). A different random starting seed number was used for each of the three trials.

Bayesian analysis was carried out with the software program MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003). I implemented the site-specific gamma distribution and allowed for invariant sites. Cytb and CR regions were partitioned separately, with the Cytb coding region further partitioned by codon position. I ran 4 Markov-chains, for 10 million generations, with a sampling frequency of every 1,000th generation. The first 1,000 trees were discarded as "burnin" and a majority rule consensus tree was created with the remaining trees. Nodal support was calculated for tree topologies using clade posterior probabilities (PP) estimated with MRBAYES 3.1.2 (Ronquist and Huelsenbeck 2003).

Divergence Time Estimation

I approximated the date of molecular divergence of major clades within the intraspecific phylogeny to allow for more detailed interpretation of the phylogeny in a historical biogeographic context. To infer that Pleistocene climate change events contributed to the divergence of clades, a time estimate is required. Fossil marsh rice rats have been found in Florida and Georgia that date from the early Sangamonian to Recent (Ray 1967, Webb 1974). Fossil remains have also been uncovered in Iowa, Illinois, Indiana, Ohio, West Virginia, and Pennsylvania, farther north than the marsh rice rat's present range (Richards 1979). An extinct subspecies *O. p. fossilis* was discovered in Texas and dates from the Kansan glacial and Sangamonian interglacial periods (Dalquest 1962, 1965). These fossils and their approximate dates, suggest that the marsh rice rat has inhabited North America since the Pleistocene (1.8 million bp – 10,000 bp). From the fossils found in Florida and Georgia, I estimated that the marsh rice rat has inhabited the southeastern United States since at least 125,000 to 75,000 years ago, the time span of the Sangamonian interglacial period. The marsh rice rat occurred in Texas even earlier since the Kansan glacial period preceded the Sangamonian.

I first used a likelihood ratio test to determine the presence of substitution rate homogeneity across the phylogenetic tree. Likelihood scores were calculated in PAUP v.4.0b10 (Swofford 2003) using the previously fitted GTR + I + G model of evolution in the absence of a molecular clock and when a molecular clock is enforced. The Cytb and CR data sets were analyzed separately since the rate of evolution may be different for each mtDNA region. The CR, because it it non-coding, evolves more quickly than Cytb which does code for proteins (Bellinvia 2004, Greenberg et al. 1983; see below). I then

compared the likelihood scores of the clock and non-clock trees using a chi-square test (df = n (number of taxa) – 2). If the difference is not significant, the null hypothesis of rate constancy will fail to be rejected and a molecular clock can be used to date nodes on the tree. If the difference in likelihood scores is significant then the molecular clock hypothesis of evolution must be rejected. If evolutionary rates vary across the tree, a relaxed molecular clock model, which allows for differing rates of evolution on different tree branches, can be implemented to estimate dates of molecular divergence. Coues' rice rat from Texas was used as the outgroup in these analyses.

The approximate date of molecular divergence of major clades within the intraspecific phylogenetic tree was calculated using the software BEAST 1.4.6 (Drummond and Rambaut 2007). This program has the capability of dating nodes on the phylogenetic tree using either a strict molecular clock or a relaxed molecular clock. BEAST uses an uncorrelated relaxed clock model, meaning there is no a priori correlation between a lineage's substitution rate and that of its ancestor (Drummond et al. 2006 and 2007). The rate for each branch can be estimated either from a lognormal distribution or exponential distribution. The authors recommend using the uncorrelated relaxed lognormal clock (Drummond et al. 2007). A relaxed molecular clock may be a more realistic than a strict clock model and the use of relaxed models has become more common for estimating divergence dates (Drummond et al. 2006, Wertheim et al. 2010).

Nucleotide substitution rates within rodents have been found to be faster than in other mammals. Li et al (1990) suggested rates of molecular evolution in rodents are at least 1.5 times higher than in other mammals. A rate of 0.023 substitutions per site per million years (my) was used as the rate of evolution for the Cytb gene. This substitution

rate was calibrated from the fossil record for the subfamily Sigmodontinae of which the genus *Oryzomys* is a part (Bonvicino et al. 2009, Smith and Patton 1993). The CR, due to its non-coding nature, is highly variable and may evolve 10 times faster than the rest of the mitochondrial genome (Bellinvia 2004, Greenberg et al. 1983). I estimated a substitution rate of 0.23 substitutions per site per my as the rate of evolution for the CR (0.023 substitutions per site per my for Cytb x 10 = 0.23). These previously determined substitution rates were used to calibrate the Cytb and CR trees in this divergence date estimation. However, the mitochondrial CR may be less useful for divergence dating because the substitution rate across this region is not constant; there are three different domains in the CR all of which evolve at varying rates (Sbisà et al. 1997). This variation may introduce a greater amount of error and lead to larger confidence intervals.

The Cytb and CR data sets were analyzed separately; similar estimates from each will give me greater confidence in the results. BEAST implements a Bayesian approach using Markov chain Monte Carlo (MCMC) simulations for divergence time estimates. Input files were created in the program BEAUTi v.1.4.8 (Drummond and Rambaut 2007). I implemented the GTR + I + G model of evolution and had the program estimate base frequencies. For Cytb, I partitioned the data set by codon position, keeping the 1st and 2nd positions in one partition and the 3rd in another, so that the substitution rate of the third position is allowed to vary more than the first two. Tuning parameters for the MCMC operators were set to auto-optimize and each MCMC chain was started from a random tree. I used a chain length of 10 million generations and sampled every 1,000 generations. Three independent runs were performed for each analysis. Results from each run were visualized in the program TRACER v1.5 and runs were combined using

LOGCOMBINER v.1.4.3 (Rambaut and Drummond 2007). The TREEANNOTATOR program in BEAST was used to summarize trees from each run with a burn-in of 10%, which discarded the first 1000 trees. I then visually examined trees with dated nodes in FIGTREE version 1.0 (Rambaut 2006).

Analyses of Genetic Divergence and Diversity among Populations

I estimated the average genetic distances among all populations and among the three hypothesized geographic regions of genetic division: southwest (Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Tennessee), southeast (Florida, Georgia, and Alabama), and northeast (Delaware, Virginia, North Carolina, and South Carolina). Because the four *O. p. palustris* samples from Mississippi and Tennessee grouped with *O. p. texensis* in all phylogenetic analyses, they were included with the southwestern group for this analysis. Genetic distances were estimated under the Kimura 2-parameter model of evolution (Kimura 1980) using MEGA4 (Tamura et al. 2000) and levels of genetic differentiation among all populations and then among the three geographic regions were inferred. Genetic distances within each group and population were also estimated. Except as noted above, groups for comparison were determined a priori based on the hypothesis of three geographic clades.

An analysis of molecular variance (AMOVA) was performed with the program ARLEQUIN v. 3.0 to quantify genetic variation at three hierarchical levels: within populations, among populations within geographic regions, and among geographic regions (Excoffier et al. 2005). If genetic diversity is geographically structured within

the marsh rice rat, I would expect most of the molecular variation to occur among geographic regions.

To visualize relationships among haplotypes, I constructed minimum spanning networks (MSN) for the Cytb and CR data. These networks show the minimum number of mutational steps between haplotypes. The MSN's were calculated in ARLEQUIN (Excoffier et al. 2005) using the algorithm of Rohlf (1973).

Regional genetic diversity was analyzed for each mtDNA region by calculating mtDNA nucleotide diversity per site (π) for each of the three geographic regions and for each population using ARLEQUIN (Excoffier et al. 2005). The π measure of nucleotide diversity is based on the mean number of pairwise differences among haplotypes in a population (Tajima 1983, 1993). Haplotype frequencies within populations, shared haplotypes among populations, and haplotype diversity (h) for each population were also determined.

The degree of isolation among populations was estimated by calculating the present level of gene flow between each population using F_{ST} measures (Weir and Cockerham 1984). Wright's F_{ST} measure (Wright 1951, 1965) can be understood as the diversity attributed to genetic differences among populations. This statistic is used to estimate how genetically differentiated two populations are. It can be used to infer levels of gene flow between populations, however an F_{ST} closer to zero, indicating no genetic differentiation between two populations, could also be attributed to the recent divergence of one population into two (Holsinger and Weir 2009).

To test for isolation by distance, I performed a Mantel test in ARLEQUIN to measure the correlation between population pairwise F_{ST} genetic distances and

geographic distances for the Cytb and CR data sets separately. Geographic distances among populations were estimated using GOOGLE EARTH (Google 2007). All distances were estimated along a land path; the distance between populations on peninsular Florida and the Gulf Coast of Texas were estimated along the coast, not across the Gulf of Mexico because marsh rice rats probably cannot cross the large distance of the Gulf between Florida and Texas.

Population Size Changes

Tajima's D is an estimate of how nucleotide diversity θ_π ($N_e\mu$, the population mutation rate) estimated from pairwise nucleotide differences differs from θ_s, nucleotide diversity based on segregating sites. Tajima's D was estimated for each population in ARLEQUIN. This statistic can be used to infer demographic history because θ_π may be more sensitive to population size (N_e) changes than θ_s (Tajima 1989a). If D = 0, the effective population size (N_e) may be stable, but if D is negative N_e may have recently changed. In a population with no polymorphism, estimates of nucleotide diversity and therefore Tajima's D will be 0, but not necessarily because the population is at neutral equilibrium. Under the neutral theory of evolution θ_π and θ_s should be about equal.

Population size changes can also be inferred by calculating a mismatch distribution. For each population with more than two individuals and more than two haplotypes, as well as for each of the three geographic regions, I plotted the observed frequencies of pairwise genetic differences among individuals. A mismatch distribution is influenced by historic demographic changes. A unimodal distribution can indicate a population expansion. Stable populations produce a multimodal or erratic distribution

and a sudden reduction in population size produces an L-shaped distribution (Slatkin and Hudson 1991, Rogers and Harpending 1992, Rogers 1995). I calculated Harpending's raggedness statistic, rg, to statistically test the goodness of fit of the observed data to a model of exponential population growth (Harpending 1994). This analysis was performed using ARLEQUIN (Excoffier et al. 2005).

Results

Intraspecific Phylogenetics

A total of 1143 base pairs were sequenced for Cytb and 1044 base pairs for the CR. After combining the two data sets and aligning all sequences in MEGA4 (Tamura et al. 2000), there were a total of 2249 positions and 133 unique haplotypes. Gap coding with FASTGAP increased the number of informative characters by 52. Nucleotide frequencies for the Cytb sequences were A = 32.8%, C = 27.1%, G = 11.7%, T = 28.4%, and for the CR sequences A = 35.4%, C = 25%, G = 10.3%, T = 29.3%. Cytb transitions were 5.1 times more common than transversions, and 2.36 times more common in the CR.

For the parsimony analysis, 495 informative characters were used to calculate 10,000 equally most-parsimonious trees (length = 1375 steps, consistency index = 0.5505, retention index = 0.9567). Maximum likelihood and Bayesian analyses estimated the same tree topology as the parsimony analysis, though relationships among haplotypes were much better resolved. The maximum likelihood and Bayesian phylogenetic analyses produced the same tree, with slight differences (see results in Chapter Two for more detail). As in the previous study of the marsh rice rat's systematics (Chapter Two),

all three phylogenetic analyses supported the presence of two separate clades, one containing eastern populations and the other containing western populations (Figure 3.2). The eastern clade (Clade A) contained samples from Colbert County Alabama, Tallapoosa County Alabama, Lewes Delaware, Everglades Florida, Franklin County Florida, Gulf County Florida, Little Pine Island Florida, Lower Florida Keys, Okeechobee County Florida, Sanibel Island Florida, St. Johns County Florida, Glynn County Georgia, Gwinnett County Georgia, Dare County North Carolina, New Hanover County North Carolina, Richland County South Carolina, Norfolk County Virginia, and Northampton County Virginia. The western clade (Clade B) contained samples from Crittenden County Arkansas, Cameron Parish Louisiana, Tamaulipas Mexico, Lee County Mississippi, Okmulgee County Oklahoma, Shelby County Tennessee, Anderson County Texas, Brazoria County Texas, Calhoun County Texas, Cameron County Texas, Freestone County Texas, Galveston County Texas, San Patricio County Texas, and Willacy County Texas. These two clades were supported by the highest bootstrap support values (100%) and highest posterior probability values (1.0). This distinct genetic disjunction is geographically located between Mississippi and Alabama, farther east than the originally proposed divide between the western subspecies *O. p. texensis* and the eastern subspecies *O. p. palustris*. The MSN's for each mtDNA region also resolved two distinct groups of eastern and western haplotypes (Figures 3.3 and 3.4).

Within each clade the branches were shallow indicating little divergence among populations and among haplotypes within populations (Figure 3.5 and 3.6). The MSN's also depicted the close relationships among haplotypes within each clade; the minimum mutational steps between haplotypes were small (Cytb 1 step to 11 steps; CR 1 step to 19

steps) compared to the number of mutational steps between eastern and western haplotypes (Cytb 59 steps; CR 92 steps). Two populations formed monophyletic groupings within the eastern clade, individuals from Sanibel Island, Florida and from the Florida Keys (BS = 100%, PP = 1.0). The Florida Keys haplotypes were closely related with haplotypes from the Florida Everglades (Miami-Dade County). In the maximum likelihood and Bayesian analyses, Florida Keys haplotypes were nested within a clade of Everglades' haplotypes. In the MSN the Florida Keys haplotypes were connected to haplotypes from the Everglades indicating a close genetic relationship. The Sanibel Island population was within a paraphyletic clade with a haplotype from Okeechobee County Florida and one haplotype from the Everglades. In the MSN, Sanibel Island haplotypes were connected to haplotypes from the Everglades. Within the western clade, individuals from extreme southeastern Texas (Cameron and Willacy counties) and northeastern Mexico (Matamoros, Tamaulipas) fell into a separate clade (BS = 100%, PP = 1.0). However, some haplotypes from Cameron County Texas clustered with other Texas haplotypes. This relationship was also depicted in the MSN.

Molecular Divergence Date Estimates

The Cytb likelihood ratio test determined that the molecular clock hypothesis could not be rejected ($\chi^2 = 109.14$, df = 111, $p = 0.5332$), but for the CR the molecular clock hypothesis was rejected ($\chi^2 = 215.16$, df = 112, $p < 0.001$). Despite the molecular clock not being rejected, I utilized both a strict clock and relaxed clock model in BEAST for Cytb. Even though the molecular clock hypothesis holds for Cytb, a relaxed clock model allowing rates to vary randomly across branches may be more realistic

(Drummond et al. 2007, Wertheim et al. 2010). I did not use a strict clock model for analysis of CR data because the molecular clock hypothesis was rejected.

For Cytb, the strict clock model dated the divergence between the eastern and western clades to 1.79 million years ago (mya) with a 95% confidence interval (95% CI) of 2.25 to 1.34 my. The time to most recent common ancestor (tmrca) for the western clade was 0.35 mya (95% CI 0.48 - 0.22 my) and for the eastern clade was 0.3 mya (95% CI 0.44 - 0.19 my), inferring that the common ancestor of the western clade may be older than the common ancestor of the eastern clade. The relaxed Cytb clock estimated the divergence between eastern and western populations to be 1.9 mya (95% CI 3.14 - 0.92 mya). The tmrca for the western clade was estimated as 0.44 mya (95% CI 0.74 - 0.21 mya), and for the eastern clade was 0.33 mya (95% CI 0.54 – 0.18 mya). The divergence date estimate between the eastern and western clades using a relaxed molecular clock for CR sequences was 0.34 mya (95% CI 0.47 - 0.22 mya), much more recent than the estimation based on Cytb data. Similarly, the tmrac for the western clade was much more recent at 0.06 mya (95% CI 0.08 - 0.04 mya) and for the eastern clade 0.05 mya (95% CI 0.07 - 0.03 mya). In this study, Cytb dates may be more plausible as this gene's rate of evolution was estimated from related Sigmodontinae species, whereas the CR estimate was based on a general observation from many different rodent species. Though it is possible for two genetic regions to diverge at different times, the Cytb gene and CR are linked on the mtDNA molecule making it unlikely that they would have such separate histories.

The divergence between the eastern and western clades occurred sometime towards the late Pliocene early Pleistocene when the Earth was falling into the first of

many Pleistocene glacial periods (Haq et al. 1977). The western clade may be older than the eastern clade by as much as 0.05 million years. The tmrca represents the haplotype that all other haplotypes in that clade are descended from, though other haplotypes probably existed within the clade at that time. The most recent common ancestor of the extant haplotypes in the western clade may be descended from an ancestor that is older than the most recent common ancestor of the extant haplotypes of the eastern clade. This does not necessarily mean that the western clade is older than the eastern clade, but that all other lineages within the western clade died out before lineages in the eastern clade.

Genetic Divergence Estimates and Geographic Structuring

Kimura 2-parameter genetic distances also supported the separation of eastern and western marsh rice rat populations. For the Cytb gene, genetic distances between populations from the eastern and western clade ranged from 5.5% to 6.3%, and for the CR between 8.8% and 9.9%. Much more genetic divergence was present between the eastern and western clades, than within either clade. In the eastern clade, genetic distances among populations were between 0.1% and 1.3% for Cytb (Table 3.1) and between 0.2% and 1.6% for the CR (Table 3.2). Within the western clade genetic divergence in the Cytb gene also ranged from 0.1% to 1.3% (Table 3.3), and for the CR between 0 and 2.4% (Table 3.4).

As would be expected from the divergence uncovered between the eastern and western clades, when populations were grouped by geographic region (northeast, southeast, and southwest) the extent of genetic divergence between the southwest region and both eastern regions was similar. Genetic distance between northeast and southeast

populations was much less than between the southwest region and both of these eastern regions. Genetic divergence among haplotypes within the southwest region was greater than genetic divergence among haplotypes within both the northeastern and southeastern populations (Table 3.5).

K2P genetic distance within individual populations was largest within the Cameron Parish Louisiana population for Cytb, but for the CR was largest within the St. Johns County Florida population. However, data for this population only included two individuals with different haplotypes, whereas the estimate of divergence within the Louisiana population was based on data from 20 individuals with 13 haplotypes. No genetic divergence was present among the haplotypes from Anderson County Texas, the Lower Florida Keys, Tamaulipas Mexico, Tallapoosa County Alabama, Gulf County Florida, and Lewes Delaware for both mitochondrial regions. However of these, only the estimates from the Florida Keys population and the Mexican population were based on 10 or more individuals. Also, there was no CR genetic divergence among haplotypes within the Sanibel Island population and the population from Willacy County Texas, both of which included more than 10 individuals (Table 3.6).

The AMOVA analysis revealed that 83.62% of the genetic variation occurs among geographic regions, 9.06% is attributable to among populations within each geographic region, and 7.32% is explained by within population variation (Table 3.7). This analysis infers that most of the genetic variation can be attributed to variation among the three geographic regions.

Genetic Diversity

There were a total of 92 Cytb haplotypes and 97 CR haplotypes from the 257 marsh rice rat individuals representing 32 populations (Table 3.8). Overall haplotype diversity for Cytb was 0.967 and for CR was 0.959. Only 13 Cytb haplotypes and 8 CR haplotypes were shared among two or more populations (Table 3.9 and 3.10). Out of 1143 nucleotide sites in the Cytb gene, 169 were variable (number of segregating sites, S = 169) and out of 1119 sites including gaps in the CR, 188 were variable (S = 188). Overall nucleotide diversity (π) was 0.032 for Cytb, and the average number of nucleotide differences between Cytb haplotypes was 36.39. Overall π was 0.049 for the CR, and the average number of nucleotide differences between CR haplotypes was 50.17. As expected, the non-coding CR was more variable than the coding Cytb gene.

Nucleotide diversity was greatest in the southwest region for both Cytb and CR haplotypes (π = 0.008 and π = 0.014 respectively; Tables 3.11 and 3.12). The northeastern populations had the lowest nucleotide diversity for both Cytb and CR (π = 0.005 and π = 0.007 respectively). Nucleotide diversity was similar for both mtDNA regions within all geographic areas ($\pi \leq 0.014$). Populations in the southeastern region had the greatest haplotype diversity for both Cytb and the CR (h = 0.899 and h = 0.901 respectively). The northeastern group had the lowest haplotype diversity for Cytb (h = 0.899), but a greater CR haplotype diversity than the southwest group (southeast h = 0.901 and southwest h = 0.894). Haplotype diversity was high for both mtDNA regions in all three geographic areas (Table 3.13).

The population from New Hanover County North Carolina had the highest genetic diversity for Cytb (π = 0.045), but the populations from Cameron Parish

Louisiana and St. Johns County Florida had the highest genetic diversity for the CR (π = 0.013). Cytb nucleotide diversity in the New Hanover County population may have been high due to sampling; some samples came from museum specimens that were collected 30 years ago introducing a time factor that could bias the results. If nucleotide diversity was higher in this population in the past, the present nucleotide diversity would appear to be greater than it currently is. Also, samples from this county were collected from multiple sites, perhaps uncovering more variation among haplotypes than that found among samples from one site. The Florida Keys population had no genetic diversity in either mtDNA region. No CR genetic diversity was found within populations from Tamaulipas Mexico, Willacy County Texas, Colbert County Alabama, and Sanibel Island Florida.

Cytb haplotype diversity was highest within the Everglades Florida population (h = 0.979). Four other populations (Calhoun County Texas, Okeechobee County Florida, Franklin County Florida, and St. Johns County Florida) had Cytb haplotype diversities of one (all haplotypes were different), but these estimates were from very low sample sizes (n = 1 to 5) compared to the Everglades sample size of 20 individuals. This same pattern and small sample size effect was also seen for CR haplotype diversity (Everglades h = 0.974; Calhoun County Texas, Franklin County and St. Johns County Florida h = 1). The Florida Keys and Sanibel Island populations had the lowest Cytb haplotype diversities (h = 0.154 and h = 0.167 respectively) even though estimates from these populations were both from more than 10 individuals. Twenty individuals were sampled from the Tamaulipas Mexico population, yet CR haplotype diversity was 0; only 1 CR haplotype was found in this population. The Florida Keys and Sanibel Island population also had

no CR haplotype diversity, as well as the Willacy County Texas population (n = 11). The Okeechobee County Florida population had no CR haplotype diversity, but this was estimated from two individuals. A larger sample size for the populations represented by few individuals may yield different results. Nucleotide diversity and haplotype diversity for each population are summarized in Tables 3.11 and 3.12.

Genetic Differentiation among Populations

In this study, F_{ST} values for Cytb and CR should be similar for each population comparison. They were estimated separately so that the higher divergence among CR haplotypes would not be masked by the lower divergence among Cytb haplotypes. Significant F_{ST} measures ($p \leq 0.05$) between eastern and western populations ranged from 0.89 (Cameron Parish Louisiana – Lewes Delaware) to 1 (Anderson County Texas – Florida Keys) in Cytb, and in the CR from 0.86 (Cameron Parish Louisiana – St. Johns County Florida and Franklin County Florida) to 1 (Anderson County Texas – Sanibel Island Florida and Florida Keys; Tamaulipas Mexico – Gwinnett County Georgia, Florida Keys, Okeechobee County Florida, Pine Island Florida, Sanibel Island Florida, New Hanover County North Carolina, Colbert County Alabama, Tallapoosa County Alabama, Lewes Delaware; Willacy County Texas – Gwinnett County Georgia, Florida Keys, Colbert County Alabama, Tallapoosa County Alabama, Gulf County Florida, Sanibel Island Florida, Lewes Delaware). These high F_{ST} values reflect the divergence between eastern and western populations. They also suggest there is little gene flow between the two clades. F_{ST} measures between the Mississippi population in the western clade and the two Alabama populations in the eastern clade were 1 (Tallapoosa County, $p = 0.34$)

and 0.97 (Colbert County, p = 0.28) for Cytb and 1 (Tallapoosa County, p = 0.34) and 0.993 (Colbert County, p = 0.24) for the CR. These populations are close geographically, but fall into different clades. The F_{ST} values were not significant probably due to the very small sample size (Mississippi, n = 1; Tallapoosa County Alabama, n = 2; Colbert County Alabama, n = 3).

Significant F_{ST} values within the eastern clade ranged from 0.28 for Cytb (between Everglades Florida and Gwinnet County Georgia and between St. Johns County Florida and New Hanover County North Carolina) to 0.98 (between the Florida Keys and Gwinnett County Georgia, between the Florida Keys and Gulf County Florida, and between the Florida Keys and Tallapoosa County Alabama; Table 3.14). For the CR, significant F_{ST} values ranged from 0.15 (between Everglades Florida and Pine Island Florida) to 1 (between Gwinnett County Georgia and Sanibel Island Florida; Table 3.15). The Sanibel Island Florida population had a significantly high F_{ST} between other populations in Florida for the Cytb gene, suggesting limited gene flow between this island population and populations on the Florida mainland. However, the CR F_{ST} was slightly lower with the Everglades population (Cytb F_{ST} = 0.68, CR F_{ST} = 0.46). From the phylogenetic analyses, the genetic similarity between the Pine Island Florida haplotypes and Everglades Florida haplotypes was evident. But based on the lower CR F_{ST} value, I cannot distinguish whether this genetic relatedness is due to gene flow or recent population divergence (Holsinger and Weir 2009). The Northampton County Virginia population, on the lower Delmarva Peninsula, and the Norfolk County Virginia population located on the shores of the southern Chesapeake Bay had higher significant F_{ST} values than may be expected for their close geographic proximity (Cytb F_{ST} = 0.67,

CR F_{ST} = 0.64). The channel of water where the Chesapeake Bay connects with the Atlantic Ocean separates these two populations. Though marsh rice rats can disperse over water, a characteristic of this channel, such as a current or depth, is preventing them from crossing.

Significant F_{ST} values within the western clade for Cytb ranged from 0.06 (between Brazoria County Texas and Cameron Parish Louisiana) to 0.96 (between Crittenden County Arkansas and Tamaulipas Mexico and between Anderson County Texas and Tamaulipas Mexico; Table 3.16). Significant F_{ST} values within the western clade for the CR ranged from 0.07 (between Brazoria County and Galveston County Texas, and between Brazoria County Texas and Cameron Parish Louisiana) to 1 (between Anderson County Texas and Tamaulipas Mexico, and between Anderson County and Willacy County Texas; Table 3.17). High significant F_{ST} values for both Cytb and CR were calculated between the Mexican population and many other western populations, except for Cameron County Texas where the F_{ST} value was 0.14. This mirrors the monophyletic clade that haplotypes from these two populations formed with haplotypes from Willacy County Texas in the phylogenetic analyses. All three of these populations are in close geographic proximity and are closely related.

No Cytb or CR haplotypes were shared between any eastern and western populations (Table 3.9 and 3.10). More haplotypes were shared among populations within the western clade than in the eastern clade. Only the Northampton County Virginia population and Lewes Delaware population shared both Cytb and CR haplotypes within the eastern clade. Interestingly, a Cytb haplotype was shared between an individual from the Everglades and an individual from Franklin County in the Florida

Panhandle. Most haplotypes within the eastern populations were unique to the population in which they occurred. Even though genetic divergence was greater within western populations than in eastern populations, there were still many more shared haplotypes among western populations. Perhaps this can be attributed to an isolation by distance effect in eastern populations, which occupy a larger geographic region than western populations.

The Mantel test for isolation by distance found a significant positive correlation between genetic distance (F_{ST}) and geographic distance among all populations. Both Cytb and CR genetic distances were highly correlated with geographic distance, indicating an isolation by distance effect (Cytb, r (correlation coefficient) = 0.535292, $p < 0.001$; CR, $r = 0.529069$, $p < 0.001$).

Population Size Changes

A significant negative Tajima's D is indicative of a recent population expansion (Tajima 1989b, Aris-Brosou and Excoffier 1996). For both Cytb and CR data, the Everglades Florida population had a significant negative Tajima's D (Tables 3.11 and Tables 3.12). The Sanibel Island Florida population also had a significant negative Tajima's D, but only for Cytb sequence data. Only one CR haplotype was found in this population, so Tajima's D was 0 for that data set.

A mismatch distribution, the number of observed frequencies of pairwise genetic differences among individuals, is influenced by historic demographic changes. A unimodal distribution can indicate a population expansion. Stable populations produce a multimodal or erratic distribution and a sudden reduction in population size produces an

L-shaped distribution (Slatkin and Hudson 1991, Rogers and Harpending 1992, Rogers 1995). The Cytb mismatch distribution for the Tamaulipas Mexico population was significant (rg =0.679, P=0.02, Figure 3.7). This distribution had an L shape indicating a recent reduction in population size. The Cytb mismatch distribution for the New Hanover County North Carolina population was multimodal, indicating a constant population size (rg = 0.684, P < 0.001, Figure 3.7). But as with genetic diversity estimates, this mismatch distribution could be biased due to the time difference among some of the samples included in this population. Freestone County Texas had a significant multimodal CR mismatch distribution (rg = 0.308, P = 0.04, Figure 3.7). This population has had a stable population size.

The populations within the southeast region as a whole may have recently experienced a population expansion as evident in the CR mismatch distribution (CR, rg = 0.031, P < 0.001, Figure 3.8). However, the Cytb mismatch distribution appeared more multimodal (Cytb, rg = 0.018, P = 0.04, Figure 3.8). Because the CR evolves more quickly than Cytb, it may reflect more recent changes to populations than the protein coding Cytb gene. A significant negative Tajima's D for both Cytb and CR also supported a recent expansion of populations in this geographic region (Table 3.13). Cytb data indicated that the northeastern populations have had a constant population size. The significant mismatch distribution was multimodal (rg = 0.107, P = 0.01, Figure 3.8). The mismatch distribution of CR sequence data indicated that the southwest region also has had a constant population size (rg = 0.019, P = 0.01, Figure 3.8).

Discussion

This study reveals that past geologic events and climatic history of the southeastern United States have shaped the genetic diversity of the marsh rice rat. Its genetic diversity is geographically structured. Although populations from the northeast do not form a separate clade from populations in the southeast as I originally predicted, strong geographic structuring exists between eastern and western populations. There is some structuring between the northeastern and southeastern populations, as indicated by the AMOVA and genetic divergence estimates, attributable to an isolation by distance effect.

As in other species of the southeastern United States, populations of the marsh rice rat do cluster into two distinct eastern and western genetic groups, but unlike other species of this region, they are not divided at the Apalachicola River. My original hypothesis that a genetic division would be present between eastern and western haplotypes was supported, but that the divide would be at the Apalachicola River was not supported. This indicates that the marsh rice rat has responded to past geologic and climatic events differently than other species occupying the same region. Individuals collected on either side of this river both grouped into the eastern clade (samples ANERR from Franklin County Florida and SJBP from Gulf County Florida). The genetic divide occurs farther west than the Apalachicola River. An individual from northwestern Alabama (MSB81543) and one from eastern Mississippi (MSB8154), which were collected less than 97 km away from each other grouped into the two different clades. Further, the genetic divergence between these two samples is large, on the order of

magnitude of species level divergence (Baker and Bradley 2006). This is a major genetic subdivision over a relatively small geographic distance.

The eastern and western clades seem to be produced by a very early Pleistocene vicariant event as opposed to a later divergence during the Pleistocene, which has been hypothesized to have shaped the genetic structure of many species in the southeastern United States (Hayes and Harrison 1992, Soltis et al. 2006, Pauly et al. 2007, Douglas et al. 2009, Jackson and Austin 2009, Fontanella and Siddall 2010). The divergence of these two clades dated to about 1.8 to 2 mya, which coincides with the end of the Pliocene and beginning of the Pleistocene, a time of major climatic deterioration (Haq et al. 1977). At the end of the Pliocene/beginning of the Pleistocene, Earth began to enter the first of many glacial periods (Emslie 1998, Martin et al. 2008). During this time, marsh rice rat populations may have became separated by unsuitable habitat, perhaps a growing expanse of drier habitat that was wetter before the glacial period began. Populations occupying the separate eastern and western refugia diverged genetically, then re-expanded their ranges after the ice sheet retreated and habitat separating the two groups once again became inhabitable. Glacial maxima are characterized by a drier climate supporting this biogeographic scenario (Bartlein et al. 1998).

The most recent common ancestor of the western clade dates to about 400,000 years ago, and for the eastern clade to about 300,000 years ago. These dates fall within the Pleistocene time frame. The western clade may be older than the eastern clade perhaps because suitable habitat existed in the west more consistently than in the east throughout the glacial cycles. This also could be a consequence of the marsh rice rat's divergence from Coues' rice rat in the region of Texas and Mexico, which is the northern

most distribution of Coues' rice rat. The marsh rice rat may have first inhabited this western region and then dispersed farther east.

As mentioned above, the east-west genetic divide in the marsh rice rat occurs further west than in most other species. Marsh rice rats may have occupied the same eastern refugia as other species, but the western refugia occupied by other species may not have supported suitable wetland habitat for the marsh rice rat. Woodrats, deer, and other species studied in the southeastern United States are not dependant on wetland habitat. Western marsh rice rat populations may have been forced farther southwest along the Gulf Coast where wetter habitat may have existed. Alternatively, the two marsh rice rat clades diverged earlier than other species; glacial refugia could have existed in different areas during different glacial periods. Further genetic studies are needed to determine the geographic limit of the two clades and whether or not they hybridize.

Gene flow is probably not present between the two clades as supported by the high F_{ST} values among eastern and western populations. Even if there was some limited gene flow between the two clades, it is not likely female-mediated. In most rodent species males disperse from their natal ranges, but females are philopatric (Greenwood 1980). Because mtDNA is inherited maternally, genetic structuring would remain. Phylogenetic analysis using nuclear genes also showed this deep genetic subdivision between eastern and western populations (Hanson et al. 2010). This evidence supports that if there is gene flow, it is minimal. Current geographic barriers potentially exist in the area that could be maintaining this genetic divergence; the Tennessee-Tombigbee waterway is a man-made connection between the Tennessee and Tombigbee Rivers and

the most southern reaches of the Appalachian Mountains extend through Tennessee into northern Alabama and Mississippi. A man-made waterway could be a barrier to dispersal if the habitat bordering the canal was uninhabitable to marsh rice rats, i.e. no wetlands. It is possible that there are few marshes along this waterway as it is not natural and is maintained by humans as a major transportation route. The higher elevation of the Appalachian Mountain foothills would also be inhospitable to the marsh rice rat. The Tombigbee River in western Alabama could be the location of the marsh rice rat's phylogeographic break as four other species in the southeastern United States exhibited a genetic discontinuity at this river: the sunfish (*Lepomis gulosus*; Bermingham and Avise 1986), water snakes (*Nerodia rhombifera* and *N. taxispilota*; Lawson 1987) and the Carolina chickadee (*Parus caroliniensis*; Gill et al. 1993). In these species, this discontinuity was attributed to a Pliocene vicariant event (Soltis et al. 2006). More detailed studies within this divergence zone will help to uncover the cause of this pattern in the marsh rice rat.

Dispersal within each clade is not limited, except for a couple of instances which will be discussed below. Moderate to low F_{ST} values were estimated among populations within each clade signifying gene flow. However, populations that are in close geographic proximity may have a low F_{ST} and have limited gene flow if these populations recently diverged. As may be expected, there is an isolation by distance effect throughout each clade, particularly between populations in the very northeastern region and the very southeastern region of the marsh rice rat's range. Few haplotypes were shared among populations especially within the eastern clade. A shared Cytb haplotype between an individual from the Everglades and an individual in the Florida Panhandle

may be evidence for high gene flow among Floridian populations, a relatively recent range expansion, or recent isolation between populations. Cytb genetic divergence among eastern populations was low (less than 1% with the exception of the Sanibel Island population) also supporting the influence of these possible evolutionary processes. In this study, sample size was low for many populations; a larger sample size may uncover different patterns or be able to distinguish the cause of high similarity among populations. Also, an increased sample size will give evidence for the presence or absence of more shared haplotypes among populations.

Habitat connectivity within each clade should play a role in the genetic relatedness among populations. Anthropogenic factors have broken up once contiguous wetlands, decreasing habitat size and altering habitat quality. This may make it more difficult for individuals to disperse among wetlands, though upland habitats may serve as sink habitat for dispersers (Kruchek 2004). These effects may be detectable within mtDNA because it evolves more quickly than nuclear genes, however anthropogenic influences may be too recent to create a detectable genetic signal.

Isolation of Floridian populations during times of higher sea level may have been incomplete. Floridian marsh rice rats are not more genetically differentiated from other populations as predicted, with two exceptions, the Florida Keys population and the Sanibel Island population. The Florida Keys population, first identified as a separate species, is geographically isolated from the mainland by more than 100 km (Goodyear 1987). F_{ST} values were significantly higher between this population and other populations in the southeast. The Florida Keys population was recently separated from the mainland, not less than 10,000 years ago, when sea levels rose after the last glacial

period. No evidence of gene flow exists, though this has not been systematically studied. Genetic diversity was low in this population and it had unique haplotypes. These haplotypes are related to haplotypes from the Everglades because of these populations recent divergence.

The Sanibel Island population is also genetically differentiated from other populations, slightly more so than the Florida Keys population. F_{ST} values between the Sanibel Island population and other populations were consistently high, suggesting limited gene flow. The amount of dispersal of Sanibel Island rice rats needs to be studied in more detail. Whether or not gene flow is occurring between this population and others is unclear because rice rats have been trapped on smaller islands around Sanibel, giving evidence that individuals may be dispersing (Emily Woods, personal communication). This also demonstrates that they may have the ability to disperse to the mainland which is not very far away. Pine Island to the north is even closer to Sanibel than to the mainland, yet F_{ST} values and genetic divergence between populations on these two islands were high, especially for populations in close geographic proximity. This suggests that the Sanibel Island rice rats are an isolated population.

The Cytb Tajima's D statistic, estimated for the Sanibel Island individuals, suggested a recent population expansion, but the mismatch distributions did not suggest any recent change to population size. A population expansion may be biologically unlikely as anthropogenic development on Sanibel Island has reduced and degraded the already limited available habitat for this population. However, major conservation and habitat management programs are now in effect on the island, protecting vital habitat for species that live there and perhaps allowing for very recent population growth in the

Sanibel Island rice rat. Nucleotide and haplotype diversity were low in this population compared to others, possibly an affect of isolation or reduced population size.

Haplotype diversity and nucleotide diversity were relatively high in the Everglades population. This population is probably very large compared to others because an extensive portion of this wetland system is protected and able to support a larger population. A larger population should contain more genetic diversity than a smaller population. Tajima's D for both Cytb and CR was significantly negative indicating a recent population expansion. Despite environmental concerns regarding the Everglades, this habitat is supporting marsh rice rat population growth. This may simply be due to the fact that it is one of the largest wetlands in the United States (Richardson 2009).

Within the western clade, populations from extreme northeastern Mexico (Tamaulipas) and southeastern Texas (Cameron and Willacy counties) formed a highly supported monophyletic clade. This area is the southernmost known distribution of the marsh rice rat. One Cameron County haplotype consistently grouped with other Texan populations, an indication that the clade diverged from Texan populations. Genetic distances also supported the close relationship between these populations; genetic divergence was relatively greater between these populations and other western populations, then the genetic divergence among them. Even though the Tamaulipas Mexico population sample consisted of 20 individuals, only two Cytb haplotypes and one CR haplotype were present, both of which were also found in the Cameron County and Willacy County Texas populations. This Mexican population had very low genetic diversity and high F_{ST} values between it and other populations in the western clade. A

possible explanation for this pattern is a recent population expansion into northeastern Mexico, causing this population to experience a founder effect. A population bottleneck could also cause the genetic patterns observed in this population and is supported by the significant Cytb mismatch distribution.

The Willacy County Texas population also had low genetic diversity, with one CR haplotype and two Cytb haplotypes from 11 sampled individuals. This population had the same two Cytb haplotypes and CR haplotypes as the Tamaulipas Mexico population. Given the very close geographic proximity of populations in this extreme southwestern clade and the same genetic profile for the Willacy County Texas and Tamaulipas Mexico populations, individuals within these three localities may be effectively part of the same population. Though not significant, F_{ST} values between the Tamaulipas Mexico and Willacy County Texas populations were 0 indicating a high level of gene flow and further supporting the complete genetic similarity between the two.

The southeast and southwest populations had more genetic diversity than populations in the northeastern region. Less diversity could indicate a recent range expansion into this region if populations were founded by relatively few individuals. Lower genetic diversity could also be a product of a relatively smaller population size compared to the southeast and southwest regions. Habitat in the northeast region may have become suitable for marsh rice rats more recently than habitat in the south. Following the last glacial period, individuals from the eastern refugia would have dispersed south, east, and west, while individuals may not have been able to disperse north right away. Habitat would have remained drier and cooler longer in the north compared to the south as the glacier retreated. Tajima's D and the mismatch distributions

signified a recent population expansion in the southeast region supporting this biogeographic explanation. However, the Cytb data supported a constant population size in the northeast region. This region also had the smallest sample size as a whole, a factor that may bias these results and not give an accurate estimation of genetic diversity and structure in this region.

Genetic diversity was greatest in the southwestern populations. Diversity could be greater in these populations if they have been more stable through time than populations in other regions. Central Texas may have been cooler, yet wetter, in the late Pleistocene, towards the end of the last glacial period (Nordt et al. 1994). Wetlands and wet meadows could have persisted along the coast providing suitable habitat for the marsh rice rat. This allowed marsh rice rat populations to remain more stable in this region compared to the southeast and southwest. A stable population would permit for more mutations to occur in the genome that would persist within the population through time. Alternatively, this could be an artifact of a larger sample size.

Three populations sampled in this study may be threatened; the Tamaulipas Mexico/Willacy County Texas population, the Sanibel Island population, and the Florida Keys population, which is listed as endangered under the United States Endangered Species Act (ESA). The Sanibel Island rice rat and Willacy County Texas population may be candidates for protection under the ESA. Further study of these populations may be essential for their future existence. All three populations harbor unique haplotypes, enhancing the genetic diversity of the species as a whole.

Conclusions

This study presents the first description of the phylogeography of the marsh rice rat. This information is necessary for understanding the evolution of this species as well as for this species' conservation and management. Eastern and western populations are separate entities and should be managed as separate evolutionary significant units (Moritz 1994, 2002). This genetic divergence was shaped by the glacial-interglacial climate changes of the late Pliocene/early Pleistocene, but what may be maintaining this differentiation remains to be explored. A detailed study of the genetic suture zone along the border of Alabama and Mississippi will determine the extent of reproductive isolation and the presence of hybridization. Uncovering the current evolutionary mechanism maintaining this differentiation will also enhance understanding of speciation and evolution in small mammals.

Both biogeographic and demographic factors, as well as evolutionary processes, have contributed to shaping the genetic structure of the marsh rice rat. This species' close association with wetlands has played a strong role in how it has responded to environmental changes. The marsh rice rat responded differently to Pliocene and Pleistocene climate changes than other species of the southeastern United States, perhaps in part due to this habitat preference. Because an assemblage of species inhabits the same habitat or region does not determine that each will respond to climactic and environmental changes in the same way. Major genetic structuring within the marsh rice rat may have been shaped earlier than other species in the southeastern United States, as most studies attributed Pleistocene barriers to gene flow as the major force shaping genetic structure. However, Soltis et al. (2006) point out that many of these studies do

not have sufficient evidence to support this hypothesis. The effects of environmental changes may also be variable through time depending on population sizes and connectivity. Species' responses to geologic and climatic events are not fixed in either space or time, and depend on demography, behavior, and natural history.

Table 3.1. Mean pairwise Kimura 2-parameter (K2P; Kimura 1980) genetic distances (%) for the mitochondrial Cytochrome b gene among *Oryzomys palustris* populations within the eastern clade. Lowest (0.1%) and highest (1.3%) values are in bold. Populations are Norfolk County Virginia (NVA), Everglades Florida (EFL), Gwinnett County Georgia (GGA), Glynn County Georgia (GLG), Lower Florida Keys (KFL), Okeechobee County Florida (OFL), Little Pine Island Florida (PFL), New Hanover County North Carolina (NNC), Colbert County Alabama (CAL), Tallapoosa County Alabama (TAL), Richland County South Carolina (RSC), Gulf County Florida (GFL), Franklin County Florida (FFL), Dare County North Carolina (DNC), Sanibel Island Florida (SFL), St. Johns County Florida (SJF), Northampton County Virginia (NHV), Lewes Delaware (LDE).

	NVA	**EFL**	**GGA**	**GLG**	**KFL**	**OFL**	**PFL**	**NNC**	**CAL**	**TAL**	**RSC**	**GFL**	**FFL**	**DNC**	**SFL**	**SJF**	**NHV**	**LDE**
NVA	0																	
EFL	0.7	0																
GGA	0.6	0.5	0															
GLG	0.5	0.6	0.4	0														
KFL	0.7	0.5	0.5	0.6	0													
OFL	0.6	0.4	0.4	0.6	0.5	0												
PFL	0.7	0.5	0.5	0.6	0.5	0.5	0											
NNC	0.5	0.7	0.6	0.5	0.8	0.7	0.8	0										
CAL	0.6	0.6	0.3	0.5	0.6	0.6	0.6	0.7	0									
TAL	0.3	0.7	0.6	0.5	0.8	0.7	0.8	0.6	0.7	0								
RSC	0.5	0.4	0.3	0.4	0.4	0.4	0.4	0.5	0.4	0.5	0							
GFL	0.6	0.5	0.4	0.4	0.5	0.5	0.5	0.6	0.4	0.6	**0.1**	0						
FFL	0.5	0.3	0.3	0.4	0.4	0.3	0.4	0.5	0.4	0.6	0.2	0.3	0					
DNC	0.2	0.6	0.5	0.4	0.7	0.6	0.7	0.5	0.6	**0.1**	0.4	0.5	0.5	0				
SFL	1.1	1.0	1.1	1.2	1.1	0.8	1.0	**1.3**	1.2	1.2	1.0	1.1	0.9	1.1	0			
SJF	0.6	0.4	0.4	0.4	0.4	0.4	0.4	6	0.4	0.6	0.3	0.4	0.2	0.5	1.0	0		
NHV	0.8	0.7	0.6	0.7	0.8	0.7	0.8	0.7	0.7	0.8	0.5	0.6	0.5	0.8	**1.3**	0.5	0	
LDE	0.7	0.7	0.5	0.6	0.7	0.7	0.7	0.6	0.6	0.8	0.4	0.5	0.5	0.7	**1.3**	0.5	0.3	0

Table 3.2. Mean pairwise Kimura 2-parameter (K2P; Kimura 1980) genetic distances (%) for the mitochondrial control region among *Oryzomys palustris* populations within the eastern clade. Lowest (0.2%) and highest (1.6%) values are in bold. Populations are Norfolk County Virginia (NVA), Everglades Florida (EFL), Gwinnett County Georgia (GGA), Glynn County Georgia (GLG), Lower Florida Keys (KFL), Okeechobee County Florida (OFL), Little Pine Island Florida (PFL), New Hanover County North Carolina (NNC), Colbert County Alabama (CAL), Tallapoosa County Alabama (TAL), Richland County South Carolina (RSC), Gulf County Florida (GFL), Franklin County Florida (FFL), Dare County North Carolina (DNC), Sanibel Island Florida (SFL), St. Johns County Florida (SJF), Northampton County Virginia (NHV), Lewes Delaware (LDE).

	NVA	**EFL**	**GGA**	**GLG**	**KFL**	**OFL**	**PFL**	**NNC**	**CAL**	**TAL**	**RSC**	**GFL**	**FFL**	**DNC**	**SFL**	**SJF**	**NHV**	**LDE**
NVA	0																	
EFL	1.5	0																
GGA	0.9	1.1	0															
GLG	0.9	1.1	**0.2**	0														
KFL	1.5	1.0	1.2	1.2	0													
OFL	1.5	0.9	1.2	1.2	0.9	0												
PFL	**1.6**	1.1	1.3	1.3	1.1	0.8	0											
NNC	0.7	1.5	0.9	0.9	1.5	1.5	1.5	0										
CAL	0.8	0.9	**0.2**	**0.2**	1.0	1.0	1.1	0.8	0									
TAL	0.7	1.4	1.2	1.2	1.5	1.4	1.5	0.9	1.0	0								
RSC	0.9	1.0	0.7	0.7	1.1	1.1	1.2	1.0	0.5	1.3	0							
GFL	0.8	1.1	0.6	0.6	1.2	1.2	1.3	0.9	0.4	1.2	0.3	0						
FFL	1.3	1.1	0.9	0.9	1.1	1.1	1.2	1.3	0.8	1.3	1.0	1.1	0					
DNC	0.5	1.4	1.1	1.1	1.5	1.4	1.5	0.8	0.9	0.3	1.2	1.1	1.3	0				
SFL	1.4	0.9	1.1	1.1	0.9	0.7	0.9	1.3	0.9	1.4	1.1	1.2	0.9	1.4	0			
SJF	1.3	1.3	0.9	0.9	1.4	1.4	1.5	1.4	0.8	1.5	0.9	0.9	1.1	1.5	1.2	0		
NHV	0.9	**1.6**	1.0	1.0	1.5	**1.6**	**1.6**	0.8	0.9	1.2	1.1	1.0	1.5	1.1	1.4	1.5	0	
LDE	0.9	1.5	1.0	1.0	**1.6**	**1.6**	1.5	0.8	0.8	1.2	1.1	1.0	1.4	1.1	1.4	1.5	0.3	0

Table 3.3. Mean pairwise Kimura 2-parameter (K2P; Kimura 1980) genetic distances (%) for the mitochondrial Cytochrome b gene among *Oryzomys palustris* populations within the western clade. Lowest (0.1%) and highest (1.3%) values are in bold. Populations are Crittenden County Arkansas (CAR), Anderson County Texas (ATX), Brazoria County Texas (BTX), Calhoun County Texas (CAT), Cameron County Texas (CMT), Freestone County Texas (FTX), Galveston County Texas (GTX), Cameron Parish Louisiana (CLA), Tamaulipas Mexico (TMX), Lee County Mississippi (LMS), Okmulgee County Oklahoma (OOK), San Patricio County Texas (STX), Shelby County Tennessee (STN), Willacy County Texas (WTX).

	CAR	**ATX**	**BTX**	**CAT**	**CMT**	**FTX**	**GTX**	**CLA**	**TMX**	**LMS**	**OOK**	**STX**	**STN**	**WTX**
CAR	0													
ATX	0.8	0												
BTX	0.6	0.6	0											
CAT	0.4	0.6	0.5	0										
CMT	1.0	1.2	1.1	1.0	0									
FTX	0.6	0.3	0.6	0.5	1.2	0								
GTX	0.7	0.5	0.5	0.5	1.1	0.6	0							
CLA	0.5	0.7	0.6	0.6	1.2	0.7	0.6	0						
TMX	1.2	1.3	1.2	1.2	0.3	**1.3**	1.2	**1.3**	0					
LMS	0.7	0.8	0.7	0.7	1.1	0.8	0.8	0.8	1.2	0				
OOK	0.4	0.5	0.5	0.5	1.0	0.6	0.6	0.6	1.1	0.7	0			
STX	0.7	0.6	0.5	0.5	1.1	0.6	0.4	0.6	1.1	0.8	0.6	0		
STN	0.5	0.5	0.5	0.5	1.0	0.5	0.5	0.6	1.1	0.7	0.4	0.6	0	
WTX	1.2	**1.3**	1.2	1.2	0.3	**1.3**	1.2	**1.3**	**0.1**	1.2	1.1	1.1	1.1	0

Table 3.4. Mean pairwise Kimura 2-parameter (K2P; Kimura 1980) genetic distances (%) for the mitochondrial control region among *Oryzomys palustris* populations within the western clade. Lowest (0) and highest (2.4%) values are in bold. Populations are Crittenden County Arkansas (CAR), Anderson County Texas (ATX), Brazoria County Texas (BTX), Calhoun County Texas (CAT), Cameron County Texas (CMT), Freestone County Texas (FTX), Galveston County Texas (GTX), Cameron Parish Louisiana (CLA), Tamaulipas Mexico (TMX), Lee County Mississippi (LMS), Okmulgee County Oklahoma (OOK), San Patricio County Texas (STX), Shelby County Tennessee (STN), Willacy County Texas (WTX).

	CAR	**ATX**	**BTX**	**CAT**	**CMT**	**FTX**	**GTX**	**CLA**	**TMX**	**LMS**	**OOK**	**STX**	**STN**	**WTX**
CAR	0													
ATX	1.2	0												
BTX	1.5	1.0	0											
CAT	1.2	0.9	1.2	0										
CMT	1.8	1.7	2.1	1.9	0									
FTX	1.1	0.5	1.2	1.0	1.8	0								
GTX	1.5	0.9	1.1	1.1	2.1	1.1	0							
CLA	1.2	0.9	1.3	1.2	2.0	1.1	1.2	0						
TMX	2.1	1.9	2.3	2.1	0.4	2.0	2.2	2.1	0					
LMS	1.3	1.2	1.7	1.4	1.8	1.4	1.6	1.5	2.0	0				
OOK	1.0	0.7	1.3	1.1	1.7	0.9	1.3	1.2	1.9	1.3	0			
STX	1.6	1.0	1.1	1.1	2.2	1.1	0.7	1.2	**2.4**	1.8	1.4	0		
STN	1.1	0.5	1.1	1.0	1.7	0.8	1.1	1.1	1.8	1.3	0.8	1.2	0	
WTX	2.1	1.9	2.3	2.1	0.4	2.0	2.2	2.1	**0**	2.0	1.9	**2.4**	1.8	0

Table 3.5. Mean pairwise Kimura 2-parameter (K2P; Kimura 1980) genetic distances (%) and standard error between and within geographic regions for the mitochondrial cytochrome b gene and control region: Southwest = Texas, Oklahoma, Arkansas, Louisiana, Mississippi, Tennessee, Southeast = Florida, Georgia, and Alabama, and Northeast = Delaware, Virginia, North Carolina, and South Carolina.

Comparison	**Cytochrome b**	**Control Region**
Northeast and Southeast	0.8 ± 0.1	1.4 ± 0.2
Northeast and Southwest	5.9 ± 0.7	9.3 ± 0.9
Southeast and Southwest	6 ± 0.7	9.5 ± 1
Within Northeast	0.5 ± 0.1	0.7 ± 0.2
Within Southeast	0.6 ± 0.1	1 ± 0.1
Within Southwest	0.8 ± 0.1	1.5 ± 0.2

Table 3.6. Mean pairwise Kimura 2-parameter (K2P; Kimura 1980) genetic distances (%) and standard error within populations of *Oryzomys palustris* with more than one individual. Numbers of individuals within populations are in parentheses.

Population	Cytochrome b	Control Region
Norfolk County VA (18)	0.3 ± 0.1	0.3 ± 0.1
Everglades FL (20)	0.5 ± 0.1	1 ± 0.1
Anderson County TX (2)	0	0
Brazoria County TX (19)	0.5 ± 0.1	1.2 ± 0.2
Calhoun County TX (5)	0.5 ± 0.2	1.2 ± 0.2
Cameron County TX (17)	0.4 ± 0.1	0.6 ± 0.1
Freestone County TX (7)	0.5 ± 0.1	0.9 ± 0.2
Gwinnett County GA (2)	0	0.4 ± 0.2
Galveston County TX (20)	0.4 ± 0.1	0.8 ± 0.2
Lower Keys FL (13)	0	0
Okeechobee County FL (2)	0.4 ± 0.2	0.8 ± 0.3
Cameron Parish LA (20)	0.7 ± 0.1	1.2 ± 0.2
Little Pine Island FL (8)	0.4 ± 0.1	0.7 ± 0.2
Tamaulipas MX (20)	0	0
New Hanover County NC (16)	0.5 ± 0.1	0.7 ± 0.2
Colbert County AL (3)	0.2 ± 0.1	0.1 ± 0.1
Tallapoosa County AL (2)	0	0
Okmulgee County OK (4)	0.5 ± 0.2	1 ± 0.3
Gulf County FL (2)	0	0
Franklin County FL (3)	0.2 ± 0.1	1 ± 0.3
San Patricio County TX (14)	0.2 ± 0.1	0.4 ± 0.1
Sanibel Island FL (12)	0.1 ± 0	0
St. Johns County FL (2)	0.4 ± 0.2	1.3 ± 0.3
Shelby County TN (3)	0.5 ± 0.2	0.8 ± 0.2
Northampton County VA (5)	0.3 ± 0.1	0.3 ± 0.1
Willacy County TX (11)	0.1 ± 0.1	0
Lewes DE (2)	0	0

Table 3.7. Analysis of molecular variance (AMOVA) for the combined mitochondrial Cytochrome b gene and control region sequence data from *Oryzomys palustris*.

Source of Variation	df	Sum of Squares	Variance Components	Percent Variation
Among geographic regions	2	9289.602	60.64438 (Va)	83.62
Among populations Within regions	29	1568.672	6.56957 (Vb)	9.06
Within populations	225	1193.874	5.30610 (Vc)	7.32
Total	256	125052.148	0.50500	

Table 3.8. Number of Cytochrome b and control region haplotypes within each *Oryzomys palustris* population. Each data set contains 257 individuals. Numbers of individuals in populations are given in parentheses after locality names. There were a total of 92 unique Cytb haplotypes and 97 unique CR haplotypes. Some haplotypes were found in more than one population (Cytb 13, CR 8; see also Table 3.9 and 3.10). See Figure 3.1 for geographic location of each population.

Population	Cytochrome b	Control Region
Everglades FL (20)	17	16
Norfolk County VA (18)	4	4
Crittenden County AR (1)	1	1
Anderson County TX (2)	1	1
Freestone County TX (7)	4	4
Brazoria County TX (19)	11	13
Calhoun County TX (5)	5	5
Galveston County TX (20)	7	9
Cameron Parish LA (20)	12	13
Cameron County TX (17)	3	3
Tamaulipas MX (20)	2	1
Willacy County TX (11)	2	1
Gwinnet County GA (2)	1	1
Glynn County GA (1)	1	1
Lower Keys FL (13)	2	2
Okeechobee County FL (2)	2	2
Little Pine Island FL (8)	3	3
New Hanover County NC (15)	7	6
Lee County MS (1)	1	1
Colbert County AL (3)	2	2
Tallapoosa County AL (2)	1	1
Richland County SC (1)	1	1
Okmulgee County OK (4)	2	2
Gulf County FL (2)	1	1
Franklin County FL (3)	3	3
Dare County NC (1)	1	1
San Patricio County TX (14)	3	2
Sanibel Island FL (12)	2	1
St. Johns County FL (2)	2	2
Shelby County TN (3)	2	2
Northampton County VA (5)	2	2
Lewes DE (2)	1	1

Table 3.9. Cytochrome b haplotypes shared among populations of *Oryzomys palustris*. Each column represents an individual haplotype. Values in columns are the number of individuals in each population with that haplotype.

Population	Hap19	Hap22	Hap23	Hap24	Hap28	Hap30	Hap33
Everglades FL	1						
Franklin County FL	1						
Crittenden County AR		1					
Brazoria County TX		1		1	3	1	1
Calhoun County TX		1		1			
San Patricio County TX		2					
Anderson County TX			2				
Freestone County TX			3				
Galveston County TX					1	2	
Calhoun County TX		1					
Cameron Parish LA							1

Population	Hap37	Hap 39	Hap40	Hap68	Hap78	Hap92
Freestone County TX			2			
Cameron Parish LA			2			
Tamaulipas MX	17	3				
Cameron County TX	12	2				
Willacy County TX	8	3				
New Hanover County NC				2		
Dare County NC				1		
Okmulgee County OK					2	
Shelby County TN					2	
Northampton County VA						2
Lewes DE						2

Table 3.10. Control region haplotypes shared among populations of *Oryzomys palustris*. Each column represents an individual haplotype. Values in columns are the number of individuals in each population with that haplotype.

Population	Hap22	Hap23	Hap25	Hap27	Hap34	Hap36	Hap40	Hap98
Anderson County TX	2							
Freestone County TX	3							
Cameron Parish LA	2				1			
Brazoria County TX		2	2	3	1			
Calhoun County TX		1				1		
Galveston County TX			3	1		2		
San Patricio County TX						2		
Tamaulipas MX							20	
Cameron County TX							13	
Willacy County TX							11	
Northampton County Virginia								2
Lewes DE								2

Table 3.11. Nucleotide diversity (π), haplotype diversity (h), and Tajima's D-statistic based on the Cytochrome b gene. For Tajima's D statistic, an asterisk indicates a significant value. Only populations with more than one individual and more than one haplotype are included. Numbers of individuals in populations are given in parentheses.

Population	Nucleotide Diversity (π)	Haplotype Diversity (h)	Tajima's D-Statistic
Norfolk County VA (18)	0.003	0.634	- 0.834
Everglades FL (20)	0.005	0.979	- 1.911* (P = 0.014)
Freestone County TX (7)	0.005	0.810	0.459
Brazoria County TX (19)	0.005	0.906	- 0.441
Calhoun County TX (5)	0.005	1	0.436
Galveston County TX (20)	0.004	0.816	- 0.496
Cameron Parish LA (20)	0.007	0.947	- 0.077
Cameron County TX (17)	0.004	0.485	0.026
Tamaulipas MX (20)	0.001	0.268	- 0.112
Willacy County TX (11)	0.001	0.436	0.850
Lower Keys FL (13)	0	0.154	- 1.149
Okeechobee County FL (2)	0.004	1	0
Little Pine Island FL (8)	0.004	0.607	- 0.372
New Hanover County NC (15)	0.045	0.8	0.589
Colbert County AL (3)	0.002	0.667	0
Okmulgee County OK (4)	0.005	0.667	2.198
Franklin County FL (3)	0.002	1	0
San Patricio County TX (14)	0.002	0.385	- 0.655
Sanibel Island FL (12)	0.001	0.167	- 1.894* (P=0.009)
St. Johns County FL (2)	0.004	1	0
Shelby County TN (3)	0.005	0.667	0
Northampton County VA (5)	0.003	0.6	1.686

Table 3.12. Nucleotide diversity (π), haplotype diversity (h), and Tajima's D-statistic based on the mitochondrial control region. For Tajima's D-statistic, an asterisk indicates a significant value. Only populations with more than one individual and more than one haplotype are included.

Population	Nucleotide Diversity (π)	Haplotype Diversity (h)	Tajima's D-Statistic
Norfolk County VA (18)	0.004	0.634	-0.91
Everglades FL (20)	0.012	0.974	-1.679* (P=0.015)
Freestone County TX (7)	0.009	0.810	0.043
Brazoria County TX (19)	0.012	0.960	0.544
Calhoun County TX (5)	0.012	1	0.646
Galveston County TX (20)	0.009	0.916	0.052
Cameron Parish LA (20)	0.013	0.953	0.344
Cameron County TX (17)	0.006	0.404	0.053
Tamaulipas MX (20)	0	0	0
Willacy County TX (11)	0	0	0
Lower Keys FL (13)	0	0	-1.149
Okeechobee County FL (2)	0.009	0	0
Little Pine Island FL (8)	0.008	0.607	-0.072
New Hanover County NC (15)	0.008	0.808	0.122
Colbert County AL (3)	0	0.667	0
Okmulgee County OK (4)	0.1	0.667	0
Franklin County FL (3)	0.012	1	2.259
San Patricio County TX (14)	0.004	0.264	-0.665
Sanibel Island FL (12)	0	0	0
St. Johns County FL (2)	0.013	1	0
Shelby County TN (3)	0.009	0.667	0
Northampton County VA (5)	0.003	0.6	1.686

Table 3.13. Nucleotide diversity (π), haplotype diversity, and Tajima's D-statistic for each geographic region based on mitochondrial Cytochrome b (Cytb) and control region (CR) sequence data. The Cytb values are to the left and the CR values are to the right. Regions are: northeast (Delaware, Virginia, North Carolina, South Carolina), southeast (Georgia, Alabama, Florida), and southwest (Mississippi, Tennessee, Arkansas, Oklahoma, Louisiana, Texas, Mexico). For Tajima's D statistic an asterisk indicates significant at the $P < 0.05$.

Region	Nucleotide Diversity (π)	Haplotype Diversity (h)	Tajima's D-Statistic
Northeast	0.005/0.007	0.899/0.901	-0.722 (P=0.26)/-0.581 (P=0.325)
Southeast	0.006/0.009	0.942/0.937	-1.703* (P=0.014)/-1.614*(P=0.028)
Southwest	0.008/0.014	0.917/0.894	-0.771 (P=0.253)/0.031 (P=0.600)

Table 3.14. Pairwise F_{ST} among eastern populations based on the mitochondrial Cytochrome b gene. An asterisk indicates a significant value ($p \leq 0.05$). Populations are Norfolk County Virginia (NVA), Everglades Florida (EFL), Gwinnett County Georgia (GGA), Glynn County Georgia (GLG), Lower Florida Keys (KFL), Okeechobee County Florida (OFL), Little Pine Island Florida (PFL), New Hanover County North Carolina (NNC), Colbert County Alabama (CAL), Tallapoosa County Alabama (TAL), Richland County South Carolina (RSC), Gulf County Florida (GFL), Franklin County Florida (FFL), Dare County North Carolina (DNC), Sanibel Island Florida (SFL), St. Johns County Florida (SJF), Northampton County Virginia (NHV), Lewes Delaware (LDE). Negative values should be interpreted as 0.

	NVA	EFL	GGA	GLG	KFL	OFL	PFL	NNC	CAL	TAL	RSC	GFL	FFL	DNC	SFL	SJF	NHV	LDE
NVA	0.00																	
EFL	0.46*	0.00																
GGA	0.60	0.28*	0.00															
GLG	0.48	0.22	1.00	0.00														
KFL	0.79*	0.49*	0.98*	0.98	0.00													
OFL	0.53*	-0.09	0.44	0.23	0.89*	0.00												
PFL	0.59*	0.1*	0.44*	0.39	0.70*	0.15	0.00											
NNC	0.35*	0.38*	0.39	0.05	0.67*	0.36*	0.46*	0.00										
CAL	0.62*	0.35*	0.57	0.67	0.94*	0.53	0.49*	0.44*	0.00									
TAL	0.38	0.47*	1.00	1.00	0.98*	0.67	0.61*	0.33*	0.84	0.00								
RSC	0.46	-0.09	1.00	1.00	0.97	-0.11	0.14	0.11	0.50	1.00	0.00							
GFL	0.60*	0.3*	1.00	1.00	0.98*	0.55	0.44	0.39*	0.74	1.00	1.00	0.00						
FFL	0.52*	-0.05	0.61	0.54	0.89*	0.10	0.14	0.31*	0.54	0.79	0.14	0.61	0.00					
DNC	-0.04	0.29	1.00	1.00	0.98	0.23	0.46	0.01	0.71	1.00	1.00	1.00	0.63	0.00				
SFL	0.83*	0.68*	0.93*	0.93	0.95	0.84*	0.80*	0.77*	0.91*	0.93*	0.91	0.93*	0.89*	0.92	0.00			
SJF	0.51*	-0.06	0.50	0.20	0.90	0.00	0.16	0.28*	0.45	0.71	-0.33	0.50	-0.04	0.33	0.88	0.00		
NHV	0.67*	0.45*	0.66*	0.61	0.89	0.57*	0.56*	0.40*	0.65*	0.76*	0.46	0.66*	0.56*	0.65	0.89	0.46	0.00	
LDE	0.69*	0.44*	1.00	1.00	0.98	0.67	0.57*	0.39*	0.81	1.00	1.00	1.00	0.76	1.00	0.94	0.67	0.29	0.00

Table 3.15. Pairwise F_{ST} among eastern populations based on the mitochondrial control region. An asterisk indicates a significant value ($p \leq 0.05$). Populations are Norfolk County Virginia (NVA), Everglades Florida (EFL), Gwinnett County Georgia (GGA), Glynn County Georgia (GLG), Lower Florida Keys (KFL), Okeechobee County Florida (OFL), Little Pine Island Florida (PFL), New Hanover County North Carolina (NNC), Colbert County Alabama (CAL), Tallapoosa County Alabama (TAL), Richland County South Carolina (RSC), Gulf County Florida (GFL), Franklin County Florida (FFL), Dare County North Carolina (DNC), Sanibel Island Florida (SFL), St. Johns County Florida (SJF), Northampton County Virginia (NHV), Lewes Delaware (LDE). Negative values should be interpreted as 0.

	NVA	EFL	GGA	GLG	KFL	OFL	PFL	NNC	CAL	TAL	RSC	GFL	FFL	DNC	SFL	SJF	NHV	LDE
NVA	0.00																	
EFL	0.49*	0.00																
GGA	0.66*	0.21*	0.00															
GLG	0.62	-0.01	1.00	0.00														
KFL	0.87*	0.43*	0.99*	0.99	0.00													
OFL	0.73*	-0.04	0.64	0.28	0.91*	0.00												
PFL	0.69*	0.15*	0.51*	0.38	0.72*	0.03	0.00											
NNC	0.28*	0.37*	0.35*	0.14	0.74*	0.49*	0.52*	0.00										
CAL	0.61*	0.18*	0.81	0.71	0.98*	0.67	0.50*	0.28*	0.00									
TAL	0.59*	0.38*	1.00	1.00	0.99*	0.71	0.59*	0.38*	0.96	0.00								
RSC	0.58	-0.06	1.00	1.00	0.99	0.22	0.35	0.27	0.88	1.00	0.00							
GFL	0.60*	0.23*	1.00	1.00	0.99*	0.64	0.52*	0.35*	0.90	1.00	1.00	0.00						
FFL	0.63*	0.08	0.21	-0.19	0.82*	0.10	0.32*	0.38*	0.31	0.48	-0.09	0.34	0.00					
DNC	0.45	0.25	1.00	1.00	0.99	0.42	0.50	0.12	0.94	1.00	1.00	1.00	0.17	0.00				
SFL	0.87*	0.46*	1.00*	1.00	0.99*	0.91*	0.74*	0.74*	0.99*	1.00	1.00	1.00*	0.82*	1.00	0.00			
SJF	0.65*	0.11	0.30	-0.56	0.91*	0.23	0.41*	0.39*	0.33	0.58	-0.47	0.26	-0.04	0.15	0.91	0.00		
NHV	0.64*	0.46*	0.78*	0.72	0.95*	0.74*	0.63*	0.31*	0.76*	0.83	0.74	0.78*	0.59*	0.76	0.95*	0.65	0.00	
LDE	0.65*	0.40*	1.00	1.00	0.99*	0.73	0.58*	0.25	0.95	1.00	1.00	1.00	0.47	1.00	1.00*	0.55	0.29	0.00

Table 3.16. Pairwise F_{ST} among western populations based on the mitochondrial Cytochrome b gene. An asterisk indicates a significant value ($p \leq 0.05$). Populations are Crittenden County Arkansas (CAR), Anderson County Texas (ATX), Brazoria County Texas (BTX), Calhoun County Texas (CAT), Cameron County Texas (CMT), Freestone County Texas (FTX), Galveston County Texas (GTX), Cameron Parish Louisiana (CLA), Tamaulipas Mexico (TMX), Lee County Mississippi (LMS), Okmulgee County Oklahoma (OOK), San Patricio County Texas (STX), Shelby County Tennessee (STN), Willacy County Texas (WTX). Negative values should be interpreted as 0.

	CAR	ATX	FTX	BTX	CAT	GTX	CLA	CMT	TMX	WTX	LMS	OOK	STX	STN
CAR	0.00													
ATX	1.00	0.00												
FTX	0.10	-0.10	0.00											
BTX	0.16	0.31*	0.13*	0.00										
CAT	-0.22	0.37	0.05	-0.02	0.00									
GTX	0.41	0.39*	0.22*	0.05	0.12	0.00								
CLA	-0.27	0.22	0.07	0.06*	-0.06	0.14*	0.00							
CMT	0.57	0.69*	0.60*	0.59*	0.57*	0.64*	0.51*	0.00						
TMX	0.96*	0.96*	0.86*	0.78*	0.88*	0.82*	0.71*	0.11	0.00					
WTX	0.94	0.95*	0.80*	0.73*	0.82*	0.78*	0.65*	0.07	-0.02	0.00				
LMS	1.00	1.00	0.32	0.34	0.28	0.48	0.19	0.62	0.96	0.94	0.00			
OOK	-0.07	0.38	0.07	0.09	-0.01	0.26*	-0.02	0.56*	0.89*	0.84*	0.33	0.00		
STX	0.66	0.65*	0.40*	0.24*	0.32*	0.27*	0.27*	0.68*	0.89*	0.86*	0.70	0.49*	0.00	
STN	-0.06	0.29	-0.01	0.08	-0.06	0.22*	-0.07	0.55*	0.90*	0.85*	0.28	-0.21	0.47*	0.00

Table 3.17. Pairwise F_{ST} among western populations based on the mitochondrial control region. An asterisk indicates a significant value ($p \leq 0.05$). Populations are Crittenden County Arkansas (CAR), Anderson County Texas (ATX), Brazoria County Texas (BTX), Calhoun County Texas (CAT), Cameron County Texas (CMT), Freestone County Texas (FTX), Galveston County Texas (GTX), Cameron Parish Louisiana (CLA), Tamaulipas Mexico (TMX), Lee County Mississippi (LMS), Okmulgee County Oklahoma (OOK), San Patricio County Texas (STX), Shelby County Tennessee (STN), Willacy County Texas (WTX). Negative values should be interpreted as 0.

	CAR	ATX	FTX	BTX	CAT	GTX	CLA	CMT	TMX	WTX	LMS	OOK	STX	STN
CAR	0.00													
ATX	1.00	0.00												
FTX	0.27	-0.16	0.00											
BTX	0.24	0.07	0.09	0.00										
CAT	0.06	0.06	-0.01	-0.01	0.00									
GTX	0.46*	0.27	0.21*	0.07*	0.14	0.00								
CLA	0.00	0.00	0.01	0.07*	0.00	0.15*	0.00							
CMT	0.67*	0.69*	0.61*	0.56*	0.59*	0.65*	0.52*	0.00						
TMX	1.00	1.00*	0.89*	0.75*	0.89*	0.81*	0.71*	0.14*	0.00					
WTX	1.00	1.00*	0.83*	0.69*	0.83*	0.77*	0.65*	0.08	0.00	0.00				
LMS	1.00	1.00	0.39	0.32	0.24	0.50	0.20	0.68	1.00	1.00	0.00			
OOK	0.09	0.11	-0.02	0.13	-0.01	0.30*	0.03	0.59*	0.92*	0.88*	0.29	0.00		
STX	0.75	0.62*	0.43*	0.24*	0.38*	0.15*	0.29*	0.76*	0.93*	0.90*	0.77	0.57*	0.00	
STN	0.24	-0.06	-0.13	0.05	-0.04	0.23*	-0.02	0.60*	0.95*	0.91*	0.36	-0.14	0.54*	0.00

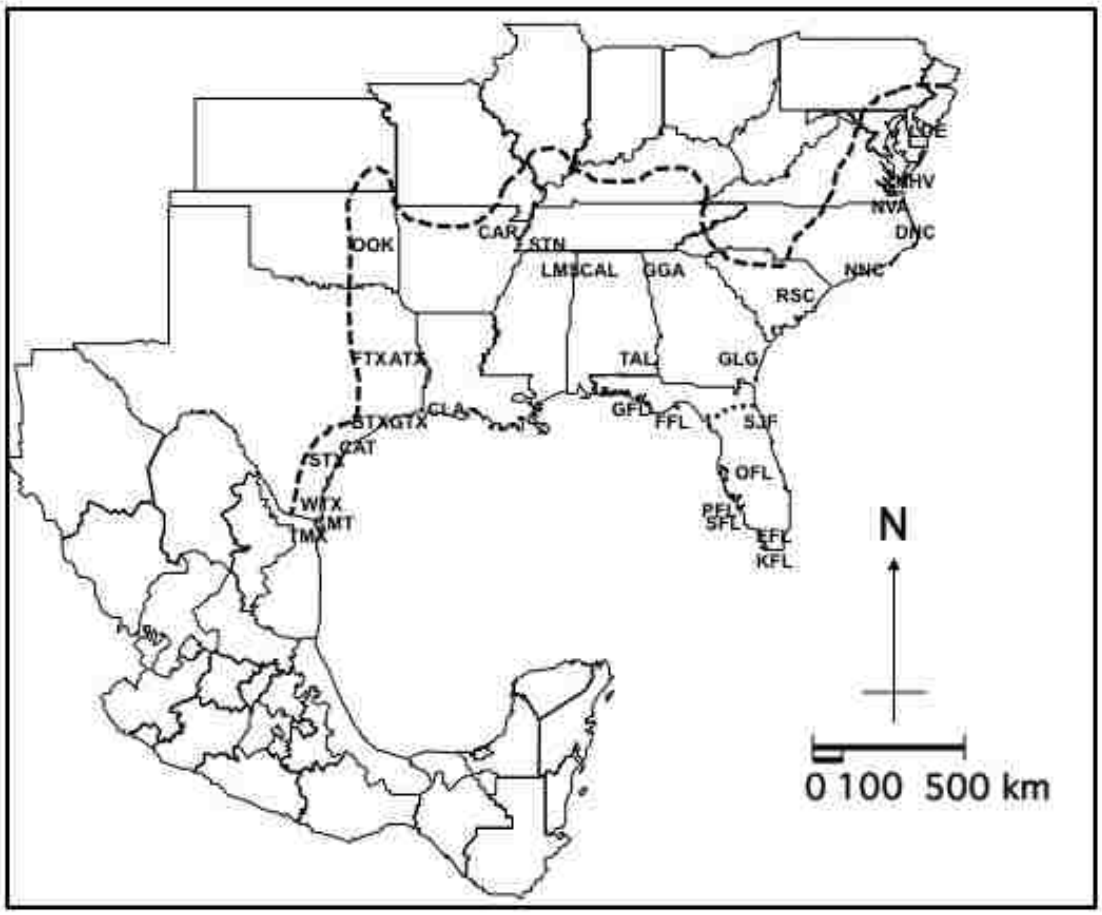


Figure 3.1. Geographic localities of *Oryzomys palustris* samples included in this study. The dotted line estimates the extent of this species' range. Population abbreviations are: Eastern - Colbert County Alabama (CAL), Tallapoosa County Alabama (TAL), Lewes Delaware (LDE), Everglades Florida (EFL), Franklin County Florida (FFL), Gulf County Florida (GFL), Little Pine Island Florida (PFL), Lower Florida Keys (KFL), Okeechobee County Florida (OFL), Sanibel Island Florida (SFL), St. Johns County Florida (SJF), Glynn County Georgia (GLG), Gwinnett County Georgia (GGA), Dare County North Carolina (DNC), New Hanover County North Carolina (NNC), Richland County South Carolina (RSC), Norfolk County Virginia (NVA), Northampton County Virginia (NHV); Western - Crittenden County Arkansas (CAR), Cameron Parish Louisiana (CLA), Tamaulipas Mexico (TMX), Lee County Mississippi (LMS), Okmulgee County Oklahoma (OOK), Shelby County Tennessee (STN), Anderson County Texas (ATX), Brazoria County Texas (BTX), Calhoun County Texas (CAT), Cameron County Texas (CMT), Freestone County Texas (FTX), Galveston County Texas (GTX), San Patricio County Texas (STX), Willacy County Texas (WTX).

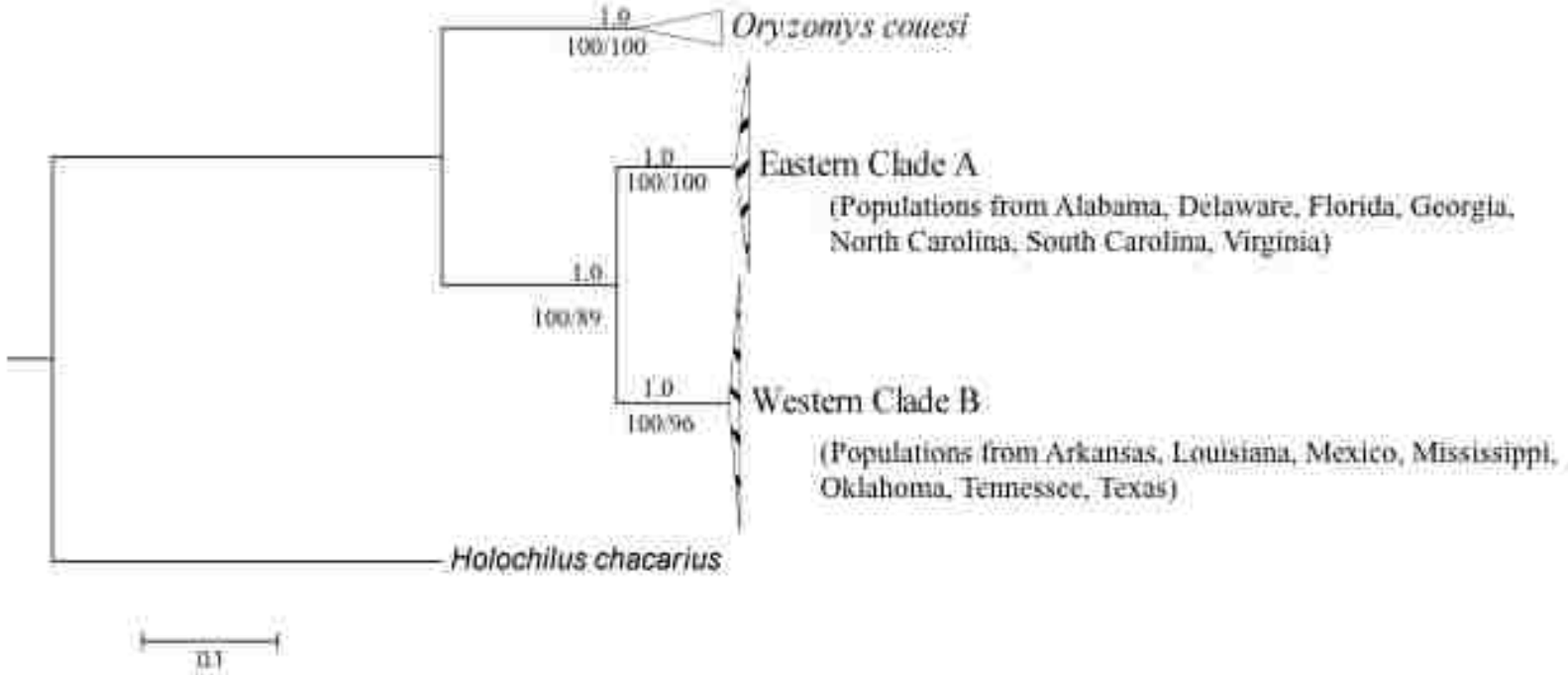


Figure 3.2. Phylogenetic tree estimated by parsimony, Bayesian, and maximum likelihood analyses of *Oryzomys palustris* mitochondrial Cytochrome b and control region sequence data. Eastern and western populations fall into two distinct clades. Posterior probabilities greater than 0.95 are given above nodes and bootstrap values (BS) greater than 60 are given below (parsimony BS/maximum likelihood BS).

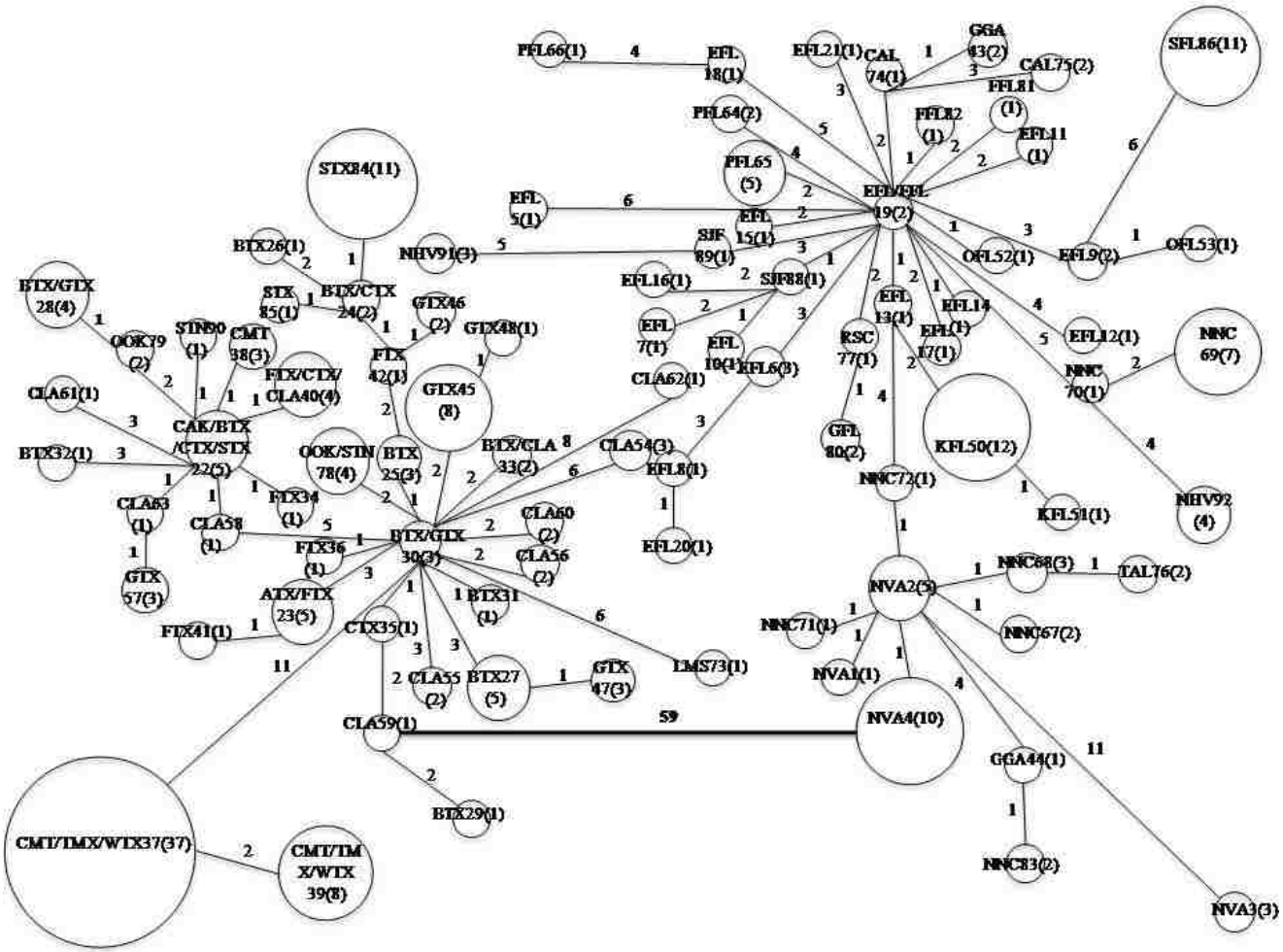


Figure 3.3. Minimum spanning network of mitochondrial Cytochrome b haplotypes sampled from *Oryzomys palustris* populations. Haplotype IDs are given in circles and numbers in parentheses refer to the number of individuals with that given haplotype. The number of mutational steps between haplotypes is given next to each line. Population abbreviations are as in Figure 3.1.

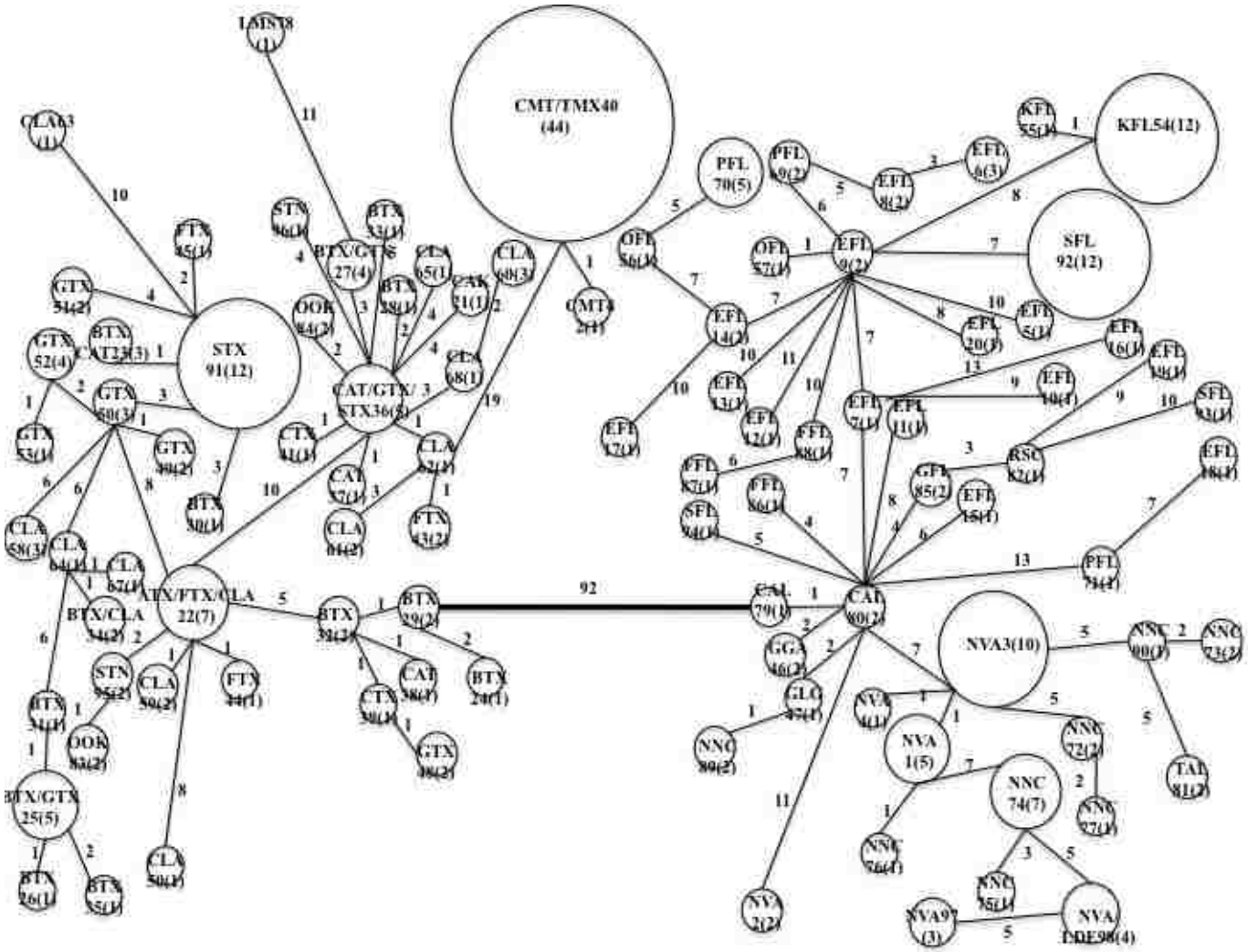


Figure 3.4. Minimum spanning network of mitochondrial control region haplotypes sampled from *Oryzomys palustris* populations. Haplotype IDs are given in circles and numbers in parentheses refer to the number of individuals with that given haplotype. The number of mutational steps between haplotypes is given next to each line. Population abbreviations are as in Figure 3.1.

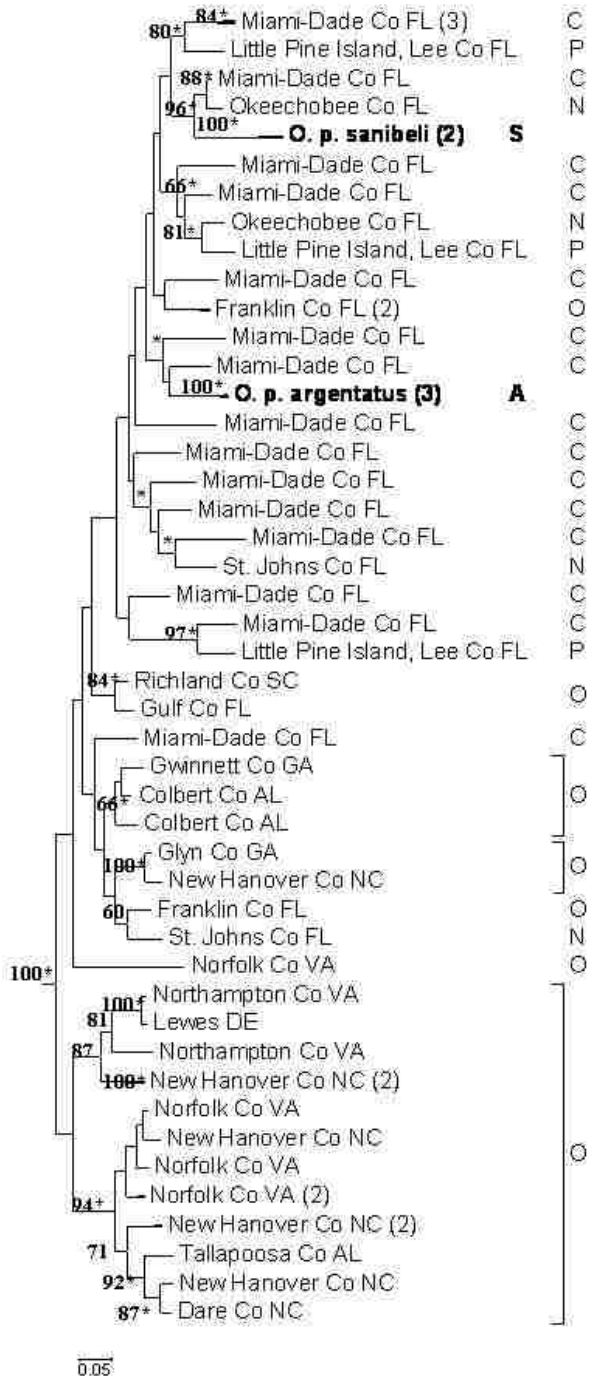


Figure 3.5. Clade A containing all eastern populations of *Oryzomys palustris* resolved by maximum likelihood and Bayesian analyses of the mitochondrial Cytochrome b gene and control region. Bootstrap support values are indicated above nodes and a posterior probability greater than 0.95 is indicated by an asterisk. This clade includes individuals of the subspecies *Oryzomys palustris palustris* (O), *O. p. natator* (N), *O. p. coloratus* (C), *O. p. planirostris* (P), *O. p. sanibeli* (S), and *Oryzomys palustris argentatus* (A). Only *O. p. sanibeli* (S) and *O. p. argentatus* (A) formed strongly supported monophyletic clades. Numbers in parentheses refer to the number of hapotypes within that collapsed clade.

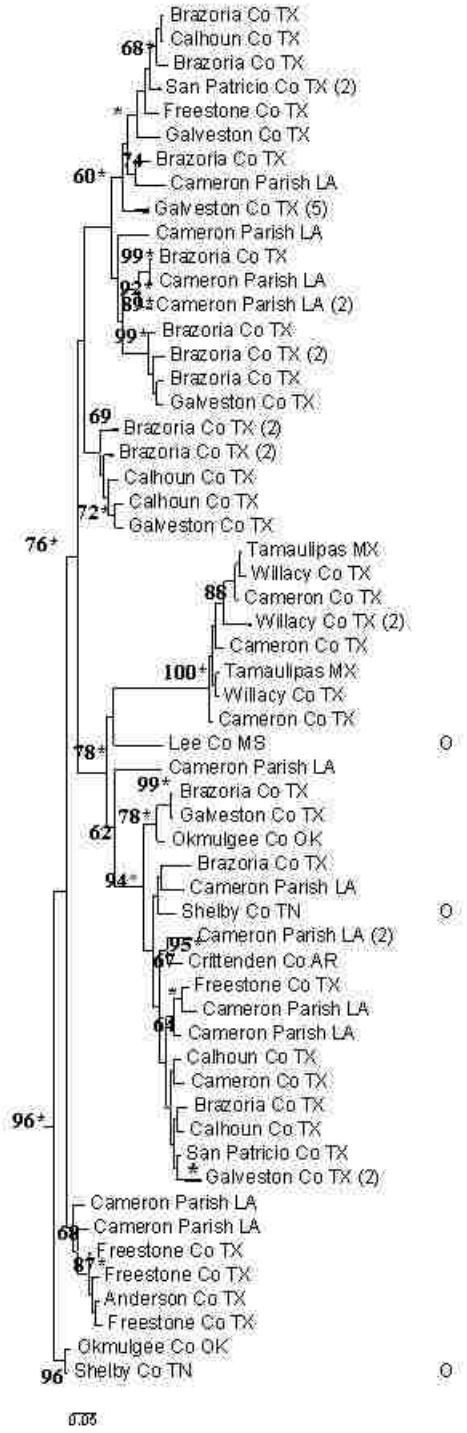


Figure 3.6. Clade B containing all western populations of *Oryzomys palustris* resolved by maximum likelihood and Bayesian analyses of the mitochondrial Cytochrome b gene and control region. Boostrap support values are indicated next to nodes. Posterior probabilities greater than 0.95 are indicated with an asterisk. This clade includes individuals assignable to *Oryzomys palutris texensis* and *O. p. palustris* individuals from Mississippi and Tennessee. Haplotypes followed by an O were originally assigned to *O.*

p. palustris. All other haplotypes are *O. p. texensis*. Populations from extreme southeastern Texas (Cameron Co. and Willacy Co. Texas) and the population from extreme northeastern Mexico consistently grouped into a monophyletic clade. However, some Cameron Co. haplotyes consistently grouped with other Texas haplotypes. Numbers in parentheses refer to the number of haplotypes within that clade.

Figure 3.7. Mismatch distributions (MD) indicating changes in population sizes. Tamaulipas Mexico Cytochrome b (Cytb) haplotypes indicated a recent population reduction (A). New Hanover County North Carolina Cytb haplotypes indicated a constant population size (B). Freestone County control region MD also signified a constant population size (C).

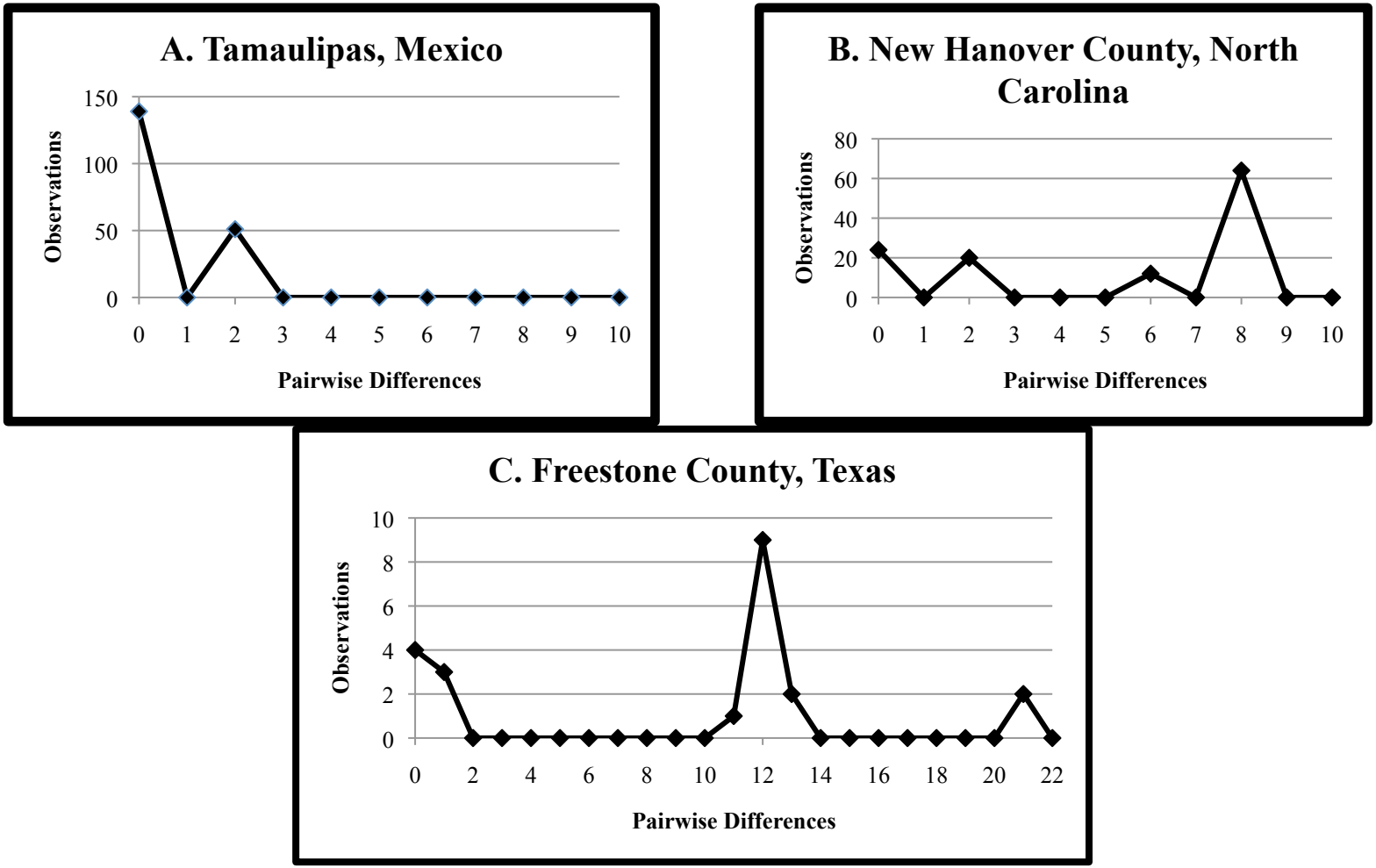
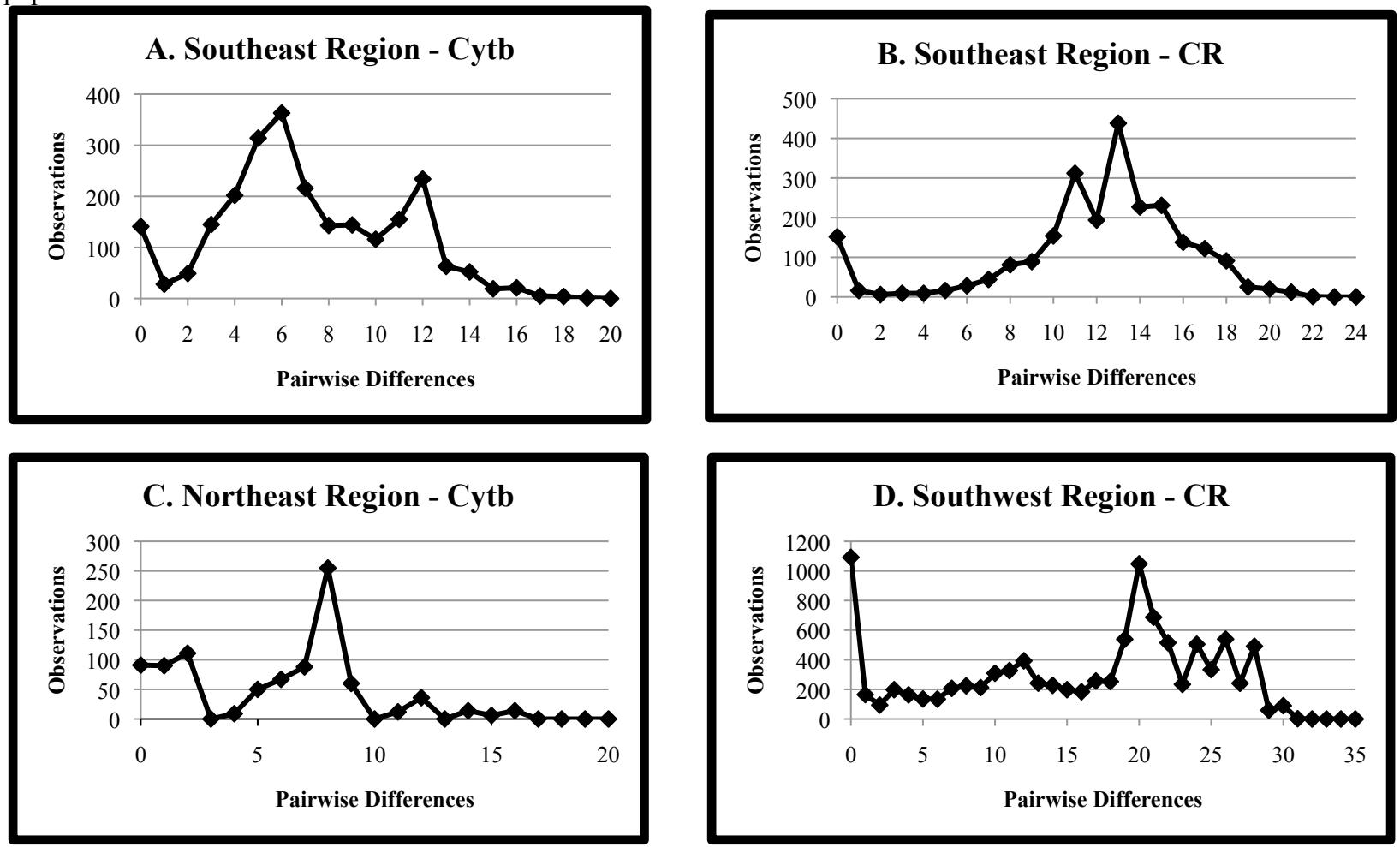


Figure 3.8. Regional mismatch distributions (MD) signifying population size changes. A mulitmodal southeastern regional Cytb MD (A) suggested a constant population size. However, the southeastern regional CR MD (B) appeared more unimodal suggesting a recent population expansion. The northeastern regional Cytb MD (C) as well as the southwestern Cytb MD (D) signified stable population sizes.



Chapter Four

Intra-population Genetic Variation and Evolutionary Processes among Populations of the Wetland Dependant Marsh Rice Rat (*Oryzomys palustris*)

Background

The evolutionary forces of natural selection, mutation, genetic drift, and gene flow all shape genetic diversity within populations and genetic structuring among populations (Avise 2000). Dispersal, the mediator of gene flow, influences many ecological and evolutionary attributes of populations (Clobert et al. 2001). Dispersal affects population demography and distributions (Bowler and Benton 2005), as well as local adaptations and speciation (Dieckmann et al. 1999). How gene flow affects a population's demography and genetic structure depends on biological attributes of the population itself, such as sex ratio and age structure, as well as environment. Habitat quality, environmental conditions, and other selective pressures will determine how a population adjusts to the changing levels of genetic diversity that gene flow can introduce (Lenormand 2002).

The effect of dispersal on genetic structure is dependant on the spatial scale under study (Peakall et al. 2003, Fontanillas et al. 2004, Gauffre et al. 2009). Dispersal may affect the genetic structure of small terrestrial mammals on a smaller geographic scale than larger mammals, because small mammals typically disperse over shorter distances and thus have reduced dispersal ability. Large scale geographic structuring has been found in small mammals primarily due to isolation by distance (Boone et al. 1999, Vega et al. 2007). Populations that are farther apart geographically are more genetically divergent because individuals mate with others that are in close proximity. Geographic structuring also is affected by sex-biased dispersal (Handley and Perrin 2007). In many

small mammal species, males typically disperse from their natal ranges, while females are philopatric (Greenwood 1980). Life history traits can strongly affect genetic diversity and structure.

Environmental influences, geologic and climatic history also will shape patterns of genetic divergence among populations. Contemporary physical barriers between populations will impede gene flow and historic barriers can still show genetic signatures within a species (Avise 2009). If gene flow is limited between two populations, genetic drift and selection may cause these populations to become genetically differentiated over time. This is especially pertinent now as historically connected populations can be separated by anthropogenic factors such as roads and development (McGregor et al. 2008, Davidson et al. 2009). Without the influx of new alleles, populations can experience a decrease in genetic diversity due to loss of heterozygosity. The population becomes inbred and loses its ability to adapt to environmental changes. Dispersal and gene flow play a key role in how populations adapt to environmental changes (Kokko and López-Sepulcre 2006).

Isolated populations have a much greater chance of becoming extinct than populations that are connected by gene flow. Genetic diversity is decreased as population size decreases and this effect is exacerbated by inbreeding, genetic drift, and reduced gene flow (Frankham 1996). Many mammal species have low levels of genetic diversity within populations that are experiencing demographic threats such as range reductions and declines in population size (Garner et al. 2005). Species that are dependent on one type of habitat may be affected more severely by anthropogenic habitat fragmentation than more generalist species. If suitable habitat for a specialist species is fragmented

with unsuitable habitat existing among more suitable fragments, dispersal and therefore gene flow may be reduced. Specialist species may not be able to disperse across unsuitable habitat. Generalist species may be more willing to cross less desirable habitat to reach high quality habitat, thereby not greatly affecting levels of gene flow within habitat generalist. Due to discontinuites in gene flow caused by habitat fragmentation, habitat specialists may exhibit more genetic structuring than habitat generalists.

The marsh rice rat (*Oryzomys palustris*) of the southeastern United States is a habitat specialist dependant on wetlands and marshes. Though this species is a habitat specialist, wetlands and rivers are not barriers to gene flow for the marsh rice rat as they may be in other small mammal species because marsh rice rats can swim and disperse over water. Marsh rice rats also move among coastal islands over open water up to 300 meters (Forys and Moncrief 1994, Loxterman et al. 1998). Therefore, gene flow likely is a strong force in shaping the genetic diversity and structure of this species. This species' intimate association with wetlands and its ability to disperse over water for longer distances than other small mammal species (Esher et al. 1978, Forys and Moncrief 1994) will be significant factors in shaping the marsh rice rat's present genetic structure and diversity. Though marsh rice rats rely on wetland habitat they do utilize upland habitat, perhaps as a sink for dispersers (Kruchek 2004). Therefore, habitat specialization of the marsh rice rat may not be a factor limiting dispersal among unconnected wetlands, even though this species is a habitat specialist. Though the marsh rice rat is dependent on wetlands, its dispersal ability may allow for more gene flow compared to other habitat specialist species which are truly restricted by their habitat requirement and dispersal behavior. Patterns of diversity in the marsh rice rat may be different than in other small

mammals of the southeastern United States, especially those with limited dispersal and specialized habitat requirements or fragemented habitat, such as the eastern woodrat (*Neotoma floridana*; Hayes and Harrison 1992), which displays limited gene flow (Castleberry et al. 2002) and the southeastern beach mouse (*Peromyscus polionotus niveiventris*), which is restricted to fragmented coastal sand dune habitat (Degner et al. 2007).

The southeastern United States is an area of great conservation concern. There are more than 90 animal taxa from this region that are listed as threatened or endangered under the United States Endangered Species Act (Avise 1996). Many of these taxa are listed due to the destruction and shrinking size of their habitat. The marsh rice rat's dependency on wetland and marsh ecosystems also makes it particularly vulnerable to future population declines. As wetland habitat is lost and degraded to human development, the marsh rice rat and other wetland species will begin to disappear. The marsh rice rat's range, habitat requirements, and life history make it an excellent indicator species for the health of wetland ecosystems throughout the southeastern United States.

In the southeastern United States wetland drainage for agriculture has changed the hydrology and connectivity among wetlands and watersheds compared to historical conditions (Blann et al. 2009). Fifty-three percent of wetlands in the lower 48 states were lost between the 1780s and 1980s mainly due to conversion of wetlands to agricultural land (Dahl 1990). This loss of habitat will have dramatic effects on this region's unique biodiversity. The southeastern United States is a transitional zone between temperate species assemblages and subtropical species creating an area of high biodiversity (Odum

2002, Blaustein 2008, Graham et al. 2010). Many wetlands have become protected areas, but these are still threatened by factors associated with anthropogenic development and global climate change (Scott et al. 2004). These threats pose a risk to the overall genetic diversity of many wetland dependent species, including the marsh rice rat.

The maintenance of a species genetic diversity is a major goal of conservation. Unique genetic variation can be protected by identifying Evolutionary Significant Units (ESU's) and management units (MU's) within a species (Ryder 1986). To conserve the most genetic diversity in a species, the genetic structure and variation of populations must be understood. This information will allow us to distinguish ESU's within a species and aid in deducing the population dynamics and natural processes that have shaped a species genetic structure. Populations with unique genetic haplotypes and those with the greatest genetic diversity can be identified and become the focus of conservation efforts.

The goal of this study was to begin to understand gene flow within the marsh rice rat and how gene flow influences the structuring of genetic diversity in this species. Understanding the genetic connectivity and degree of gene flow among populations of the marsh rice rat will aid in understanding how this species is evolving and how populations may respond to environmental changes. Mitochondrial and nuclear sequence data from the marsh rice rat have shown significant genetic divergence between eastern and western populations. This genetic disjunction exists around the Alabama-Mississippi border (Hanson et al. 2010, Chapter Three). Other vertebrate species in the southeastern United States also show east-west genetic structuring, though the genetic discontinuity in other species is farther east around the Apalachicola River (Avise 1996, Soltis et al. 2006). Whether or not gene flow is present between eastern and western populations of

the marsh rice rat is unknown, though both mitochondrial and nuclear gene sequences indicate little to no gene flow between the two groups (Hanson et al. 2010, Chapter Three). Dispersal within this species has not been extensively studied in an ecological context, so little is known about the marsh rice rat's dispersal behavior. Mitochondrial DNA supports historical gene flow among populations in both the eastern and western groups. Using microsatellite markers will allow for more fine-scale analysis of present gene flow and genetic structuring at the population level.

I hypothesized that because populations of the marsh rice rat inhabit different geographic areas within the United States, nuclear microsatellite genetic variation is geographically structured. I also originally predicted that though the marsh rice rat may be divided into eastern and western clades, gene flow might still be occurring between them because water is not a barrier to dispersal for the marsh rice rat as it may be for other species. Within the eastern and western clades, I predicted to find a significant amount of gene flow among populations.

Both historic and contemporary evolutionary processes, as well as geology and climate, influence the present genetic structure of the marsh rice rat. Populations in the southeastern region of the marsh rice rat's range will have more genetic diversity than northeastern populations because colonization of the northern region of the species' range likely happened at the end of the last glacial period when habitat became suitable farther north. Therefore, the southeastern and southwestern populations may be older than northeastern populations, and I would expect them to have more genetic diversity.

Methods

Sample Collection and DNA Extraction

I collected tissue samples from marsh rice rats throughout the species' range. Samples from 12 populations were included to incorporate all geographic areas (Figure 4.1). Between 10 and 20 individuals were sampled from each population. In this study a population is a group of individuals occupying the same sampling area and adjacent habitats or in the case of samples obtained from museums, individuals from the same county.

I obtained tissue samples by trapping individuals in Sherman live traps and cutting approximately 0.5 cm of tail tip from each animal captured using a pair of scissors. Tissue samples were stored in 1.5 ml screw cap tubes filled with a 20% DMSO (6 M NaCl) solution. Sampling methods were approved by the University of Miami Animal Care and Use Committee and followed methods approved by the American Society of Mammalogists Animal Care and Use Committee (Gannon et al. 2007). Additional samples were loaned from museum collections (tail tip, liver, or toe bone, Appendix A).

I included a total of 201 individuals of the marsh rice rat in this study. Populations analyzed were from the Everglades, Florida (Miami-Dade County, n = 20), Sanibel Island Florida (Lee County, n = 12), Lower Florida Keys (Monroe County, n = 13), New Hanover County, North Carolina (n = 19), Norfolk County Virginia (n = 17), Cameron Parish, Louisiana (n = 20), Brazoria County, Texas (n = 19), Cameron County, Texas (n = 17), Galveston County, Texas (n = 19), San Patricio County, Texas (n = 14), Willacy County, Texas (n = 11), and Tamaulipas, Mexico (n = 20). All samples and their

specific localities are listed in Appendix A. I isolated genomic DNA from tail tips and liver using a standard ethanol precipitation procedure. A DNeasy® tissue kit (Qiagen Inc., Valencia, California) was used to extract genomic DNA from museum toe bones. Some of the museum samples from the North Carolina population were collected 20 years apart, while other samples were collected within a shorter time period. For a study of genetic population structure and gene flow, ideally all samples should be from the same generation as allele frequencies typically vary over time. Samples from different generations can be tested for the absence of significant structuring between them, but a small sample size will drastically weaken this test's statistical power (Balloux and Lugon-Moulin 2002). I chose to include samples from different generations in this study because a) sample size is small for each population, so any test for temporal variation would have low statistical power, and b) this study is interested in a general interpretation of the amount of gene flow and structuring within the marsh rice rat. The bias caused by a few individual samples being from different generations is unlikely to affect the general conclusions of this study.

Microsatellite Genotyping

Nuclear microsatellite genotype data were collected for nine loci developed specifically for the marsh rice rat (Table 4.1; Wang et al. 2000). Loci were amplified via PCR using 5' end fluorescent dye-labeled primers (Applied Biosystems, Foster City, California). Forward primers were dye-labeled as in Table 4.1. Reaction volumes were 10 μl and contained 1 μl 10x PCR buffer, 3.4mM (loci AAT 03, 21, 28, 40) or 4mM $MgCl_2$ (loci AAT10, 16, 26, 60, 64), 0.1 μl Taq DNA polymerase, and 0.1mM dNTPs; see Table 4.1 for primer concentrations. Thermal profiles were: initial denaturation at

94°C (1 min), 30 cycles with denaturation at 94°C (15 s), annealing at 50 - 55°C (30 s), extension at 72°C (1 min), and a final extension at 72°C (1 min). Specific annealing temperatures for loci are given in Table 4.1. Loci with the same annealing temperatures were multiplexed and PCR products for loci with annealing temperature of 50° and 53° were co-loaded for analysis. All samples were analyzed on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Loci were scored and compiled using the software STRAND (Toonen and Hughes 2001).

Analyses of Genetic Variation within and among Populations

I tested for deviations from Hardy-Weinberg equilibrium (HWE) and genotypic linkage equilibrium using ARLEQUIN 3.1 (Excoffier et al. 2005). To determine whether low heterozygosity is a technical artifact, rather than a biological phenomenon, I checked the data with MICRO-CHECKER (Van Oosterhout et al. 2004). This software tests for the presence of null alleles, large allele dropout, and scoring errors that result in deviations from HWE. These three errors can result in overestimation of homozygosity and will increase the putative level of inbreeding within a population (Dewoody et al. 2006). MICRO-CHECKER is able to differentiate between these scoring and amplification errors and real effects such as population sub-structuring or inbreeding because these technical errors should only affect a subset of loci. A real biological effect is more likely to influence all loci and MICRO-CHECKER will warn the user if it detects this (Van Oosterhout et al. 2004).

If MICRO-CHECKER detected the presence of null alleles, their frequencies were estimated and adjusted allele frequencies were calculated (Brookfield 1996). However,

these adjusted allele frequencies could not be used in subsequent multilocus analyses because the frequencies are estimated for each population not for each individual within the population. The presence of null alleles has been shown to minimally affect the accuracy of assignment tests (Carlsson 2008). Therefore, tests of population structuring and assignment of individuals to populations, as will be performed with the program STRUCTURE (Pritchard et al. 2000, Pritchard and Wen 2003), should not be affected by null alleles. The adjusted allele frequencies can be used when calculating measures for population differentiation such as R_{ST} and F_{ST} (Van Oosterhout et al. 2006). This may not affect the outcome of the results significantly as F_{ST} has been shown to be overestimated only when populations are significantly differentiated (Chapuis and Estoup 2007). In this study adjusted allele frequencies were not used in subsequent analyses.

This genotype data was used to estimate genetic variation and diversity within each population, as well as gene flow and genetic differentiation among these populations. Though gene flow may not be happening directly due to the large geographic distance among populations in the study, gene flow may occur indirectly, with individuals dispersing among intermediate populations. DNA sequence data have shown that eastern and western marsh rice rat populations are genetically differentiated at a level comparable to differentiation between species (Hanson et al. 2010; see Chapter Two). Whether or not gene flow is happening between these two groups is unknown. This study will not be able to give conclusive evidence for the presence of gene flow between eastern and western populations because this test should be done using populations in the vicinity of the genetic divide. None of the populations used in this study are from this area.

The mean number of alleles per a microsatellite locus, observed heterozygosity (H_o) and expected heterozygosity (H_E; Simonsen et al. 1998) were calculated using the program ARLEQUIN v. 3.0 (Excoffier et al 2005). Gene flow and genetic distances between populations were calculated using the statistic F_{ST} (Wright 1951). R_{ST} was also used as an estimate of gene flow among populations (Slatkin 1995). I approximated the number of migrants (M) per generation among populations (Slatkin 1991). These three statistics were also calculated in ARLEQUIN. Slatkin's R_{ST} may be a less accurate measure in this study as simulation studies have shown that F_{ST} is a better estimator of gene flow than R_{ST} (has a larger variance than F_{ST}) when sample sizes are moderate to small ($n_s \leq 10$) and the number of loci scored is low ($n_l \leq 20$; Gaggiotti et al. 1999). To better visualize levels of gene flow, I performed a hierarchical cluster analysis on the F_{ST} and M data using the program CLUSTER v. 3.0 (de Hoon 2002). I used the program TREEVIEW to view the results of the cluster analysis and adjust the visual output (Saldanha 2004).

To describe the amount of genetic variation within each population, I estimated Nei's average gene diversity (H_E – equivalent to expected heterozygosity) and mean allelic richness corrected for sample size (El Mousadik and Petit 1996) using FSTAT 2.9.3 (Goudet 2001). F_{IS} was calculated to assess levels of inbreeding within populations. This statistic measures the correlation of genes within individuals belonging to the same subpopulation (Wright 1951). F_{IS} is an estimate of homozygosity within a population. These calculations also were carried out in ARLEQUIN.

Demographic Changes

I used the program BOTTLENECK to test for a recent reduction in effective population size (N_E) within each population (Cornuet and Luikart 1996, Piry et al. 1999). Populations that have recently experienced a bottleneck, resulting in a reduction in N_E, will show excess heterozygosity. Directly following a bottleneck, the number of alleles will be lost faster than the population's heterozygosity. The heterozygosity (H_E) of a population that has recently experienced a bottleneck will be greater than that expected at mutation-drift equilibrium (H_{eq}). Therefore, there will be a heterozygote excess in populations that have experienced a recent genetic bottleneck which can be detected in a sample of individuals from that population (Cornuet and Luikart 1996, Piry et al. 1999). I implemented the two-phase mutation model (TPM) with a probability of 95% for single step mutations (SSM) and 5% for multi-step mutations as recommended by the program for 2,000 replicates. The TPM model may be better at estimating mutations within microsatellite loci than the SSM model (Piry et al. 1999). BOTTLENECK also employs a visual method for detecting a population bottleneck. Each population's allele frequency distribution was graphed; a population that has recently experienced a bottleneck will exhibit a mode shift, while a population in mutation-drift equilibrium will have an L-shaped distribution. The New Hanover North Carolina population was not included in this analysis because as mentioned above samples from this population were collected as much as 20 years apart.

Genetic Structuring

I conducted an analysis of molecular variance (AMOVA) in ARLEQUIN to assess at which level most of the genetic diversity is found (Excoffier et al. 1992). The amount of variation attributable to within individuals, among individuals within populations, among populations, and among geographic regions was estimated. Geographic regions were Northeast (Norfolk County, Virginia; New Hanover County, North Carolina), Southeast (Everglades, Miami-Dade County, Florida; Lower Keys, Monroe County, Florida; Sanibel Island, Lee County), and Southwest (Cameron Parish, Louisiana; Brazoria County, Texas; Cameron County, Texas; Willacy County, Texas; San Patricio, County Texas; Galveston County, Texas; Tamaulipas, Mexico).

Another analysis of the degree of genetic differentiation among populations was performed using a Bayesian clustering method in the software STRUCTURE v. 2.2 (Pritchard et al. 2000, Pritchard and Wen 2003). The number of populations present in the data set (K) was estimated without geographic location data and individuals were subsequently assigned to these populations. I ran the analysis with a burn-in period of 10,000 iterations and then ran the Markov chain Monte Carlo (MCMC) simulation for 1,000,000 iterations. I ran at least five independent simulations for each K, the number of potential populations, between 1 and 15, the true number of populations plus three (Evanno et al. 2005). I implemented the program's recommended settings, using the correlated allele frequencies model and assuming admixture. This allows a proportion of an individual's ancestry to come from more than one population.

I also estimated K with the same procedure for eastern populations and western populations separately, representing the two genetically differentiated groups. Because

the mutation rate in microsatellite loci is relatively great and because back mutations can possibly generate homoplasy (as in the stepwise mutation model), structuring or lack of detected between the eastern and western populations may not be a true reflection of their genetic differentiation. An eastern individual could possibly have a genotype that clusters better with a western population due to homoplasy. Identical alleles in eastern and western populations probably do not share the same ancestor. Eastern and western populations have been divided longer than any separation that may exist within each group, as shown by mtDNA sequence data (Chapter Three). Individual assignment may be more accurate if each geographic group is analyzed separately. For the five eastern populations I ran the simulation for K = 1 – 8, and for the seven western populations for K = 1 – 10.

Pritchard et al. (2000) suggest that the best estimate of K is achieved with the maximum log probability of the data (Pritchard and Wen 2003). However, Evanno et al. (2005) found ΔK, based on the second order rate of change of the likelihood function with respect to K, was a better estimator of the true K, when K was greater than two, in many situations, for example when dispersal patterns among groups vary. Therefore, I used both the estimated log probability of the data (Pritchard et al. 2000) and ΔK calculated as in Evanno et al. (2005) to estimate the true number of populations.

Results

All nine microsatellite loci were polymorphic in each population. Across all loci in all 12 populations there were 144 alleles and the number of alleles per locus ranged from 12 (AAT16) to 18 (AAT28; Table 4.2). Twenty unique alleles were found across

populations. All loci exhibited unique alleles within populations except for AAT40. The number of unique alleles within populations ranged from one (Cameron Parish Louisiana) to five (Brazoria County Texas). The populations from Norfolk County Virginia, Cameron County Texas, Willacy County Texas, and the Florida Keys did not harbor any unique alleles (Appendix B).

Two or more loci in each population tested showed significant deviations from HWE (Table 4.3). However, Micro-Checker detected the presence of null alleles in all populations and their estimated frequencies are listed in Table 4.4. Null alleles were not detected in loci AAT16, AAT21, AAT40, or AAT60. Locus AAT64 had null alleles in most populations and this locus was the most difficult to amplify, at times having to attempt amplification three times. If this locus did not amplify after three attempts, alleles at that locus were coded as missing data. MICRO-CHECKER did not detect reduced heterozygosity across all loci within each population, so I assumed this reduced heterozygosity was due to the technical error of null alleles as opposed to a real biological phenomenon. Adjusting allele frequencies within these loci allowed them all to conform to HWE. However, within nine populations, there were loci that were not in HWE and not detected in MICRO-CHECKER as having null alleles; these I assumed to be real (Table 4.3; Florida Keys, AAT28; New Hanover County North Carolina, AAT10, AAT26; Norfolk County Virginia, AAT40; Cameron Parish Louisiana, AAT26; Brazoria County Texas, AAT64; Cameron County Texas, AAT26, AAT40; Willacy County Texas, AAT03, AAT28; San Patricio County Texas, AAT64; Galveston County Texas, AAT26). Conversely there were two instances in which MICRO-CHECKER detected null alleles in loci that exhibited no significant deviation from HWE.

All 12 populations had loci that tested significant for linkage disequilibrium (LD) at $p < 0.05$. Though its use is controversial and there are no standard conventions of when to use this correction (Moran 2003), I used a sequential Bonferroni correction for multiple comparisons to adjust the value at which p is significant ($\alpha = 0.05$; Rice 1989). After this correction, the number of significant comparisons decreased, however LD remained between loci in many populations (Table 4.5). No loci were in LD in the Willacy County Texas population. Loci AAT03 and AAT28 were in LD in all but two populations. In the two other genetic studies specifically comparing the Florida Keys rice rat population to the Everglades population, only one of them found loci AAT03 and AAT28 to be in LD in both populations (Crouse 2005, Wang et al. 2005). Because LD was not found consistently between these two loci in this study and is variable in other studies, I decided to keep both loci in subsequent analyses (Selkoe and Toonen 2006).

Genetic diversity was variable among the 12 populations (Table 4.6). The Willacy County Texas population had the least amount of genetic diversity ($H_E = 0.529$, $R = 3.908$). The Brazoria County Texas population had the highest genetic diversity ($H_E = 0.870$, $R = 8.280$). The New Hanover County North Carolina population also had one of the highest estimates of genetic diversity ($H_E = 0.863$, $R = 7.629$). A significantly positive F_{IS} value was estimated in all but three populations, Cameron County Texas, Sanibel Island Florida, and the Everglades Florida populations. These populations did not have any indication of inbreeding, though all others did. Populations with significant F_{IS} values had a heterozygote deficiency. Null alleles were not accounted for in this analysis, which could cause populations to appear to have a greater heterozygote

deficiency than they actually do (Dewoody et al. 2006). Though some level of inbreeding within these populations cannot be discounted.

F_{ST} and R_{ST} measures of genetic differentiation among populations gave similar results (Figure 4.2, Table 4.7 and 4.8). All but one F_{ST} pairwise comparison between populations were significant ($P < 0.01$). Some authors have suggested a scale for the interpretation of F_{ST}: 0 – 0.05 indicates little to no genetic variation, 0.05 – 0.15 suggests moderate differentiation, 0.15 – 0.25 indicates a large amount of genetic differentiation, and above 0.25 indicates very great genetic differentiation (Wright 1978, Hartl and Clark 1997, Balloux and Lugon-Moulin 2002). The greatest amount of genetic variation was found between the eastern Sanibel Island Florida population and the western Willacy County Texas population ($F_{ST} = 0.317$), and the least amount of genetic differentiation was estimated between the New Hanover County North Carolina population and the Everglades Florida population ($F_{ST} = 0.023$). There is less genetic differentiation between the New Hanover County North Carolina population and Everglades Florida population than there is between the Sanibel Island Florida population and the Everglades population ($F_{ST} = 0.078$), even though the former are more geographically separated. The F_{ST} estimate between the Sanibel Island population and Everglades population indicates moderate differentiation. The same trend is apparent between the Everglades population and the Florida Keys population ($F_{ST} = 0.142$). Between eastern and western populations genetic differentiation estimates ranged from very great ($F_{ST} = 0.317$) between Willacy County Texas and Sanibel Island Florida to very little ($F_{ST} = 0.024$) between the Cameron Parish Louisiana population and the New Hanover County North Carolina population. These lower values, indicating little genetic differentiation, may not be a true

reflection of the genetic differentiation between eastern and western populations due to homoplasy within these populations. Some eastern and western populations have the same alleles due to chance mutation, not because of a shared ancestry. This homoplasy may cause eastern and western populations to appear more closely related, or less genetically divergent, than they really are. R_{ST} estimates showed the same patterns of genetic differentiation as F_{ST}, though the values were greater (Table 4.8).

The absolute number of migrants (M) between eastern and western populations may not correctly estimate the actual amount of dispersal between these two groups. As with F_{ST} and R_{ST}, homoplasy may increase the apparent level of migration between eastern and western populations. M values between eastern and western populations ranged from relatively high (20.45 between the Cameron Parish Louisiana population and the New Hanover County North Carolina population) to relatively low (1.076 between Willacy County Texas and Sanibel Island Florida; Figure 4.3 and Table 4.9). A greater M was estimated between the Everglades Florida population and the New Hanover County North Carolina population (M=21.006), then between the Everglades population and the other two Floridian populations (between the Keys M=3.019, between Sanibel Island M=5.912). Within western populations relatively few migrants were estimated between the Tamaulipas Mexico population and other western populations. The largest number of migrants was estimated between the Cameron Parish Louisiana population and the Brazoria County Texas population (M = 50.112).

Demographic Changes

The Cameron Parish Louisiana population and the Brazoria County population were the only two populations that tested significant for heterozygosity excess under mutation-drift equilibrium (one-tailed Wilcoxon test for H_E excess, $P < 0.01$, $P = 0.024$ respectively). These populations may have experienced a recent reduction in N_E due to a population bottleneck. None of the populations tested showed a mode shift in their allele frequency distribution.

Genetic Structuring

Most of the genetic variation uncovered in this study was attributable to the variation within individuals and among individuals within populations (Table 4.10). The AMOVA analysis did not detect a difference in diversity levels among the three geographic regions. This result suggests that there is little variation in levels of genetic diversity among populations and among geographic regions.

The STRUCTURE software correctly determined the number of populations when the eastern and western populations were tested separately ($K = 5$, $\Delta K = 4$; $K = 7$, $\Delta K = 7$ respectively; Figures 4.4 and 4.5). However, when the simulation was run with all 12 populations, the analysis detected only nine populations ($K = 9$, $\Delta K = 4$). The method of Evanno et al. (2005) for determining the number of populations (ΔK) did not always identify the correct number of populations. There is debate over which method is more appropriate for determining the number of populations (Pritchard and Wen 2003, Evanno et al. 2005, Hubisz et al. 2009), therefore I decided to use the results from the method originally described by Pritchard et al. 2000. Individuals from Sanibel Island Florida, the

Lower Keys Florida, Norfolk County Virginia, Galveston County Texas, Cameron County Texas, Willacy County Texas, and Tamaulipas Mexico clustered together strongly (Table 4.11). No structuring was detected among individuals from the Everglades Florida population and the New Hanover County North Carolina population. All 12 individuals from the Sanibel Island Florida population were assigned to the correct population with a proportion of ancestry greater than 90%. Ten out of 13 individuals from the Florida Keys population were assigned to the correct population with a proportion of ancestry greater than 90%. Five populations (Everglades Florida, Cameron Parish Louisiana, Brazoria County Texas, San Patricio County Texas, New Hanover County North Carolina) had no individuals correctly assigned or assigned with more than 90% ancestry.

In the STRUCTURE analysis of eastern populations only, individuals from Sanibel Island Florida were correctly clustered together, with all individuals having an inferred proportion of ancestry greater than 93% (Figure 4.4). Similarly, the Florida Keys individuals grouped together with most individuals assigned with a proportion of ancestry greater than 95%. Individuals from the other three populations did not cluster together as strongly, but enough so the correct number of populations was estimated. Among the western populations, individuals from the Tamaulipas Mexico population were assigned correctly with most individuals having a proportion of ancestry between 82 and 98% (Figure 4.5). Most individuals in the Cameron County Texas population were assigned correctly and all individuals from the Willacy County Texas population were assigned correctly, with most having an inferred proportion of ancestry greater than 90%.

Individuals from other populations either grouped among each other or were assigned correctly with a low proportion of ancestry.

Discussion

This study indicates the presence of geographic structuring of genetic variation throughout the marsh rice rat's range, supporting my original hypothesis. Though the AMOVA analysis attributed the least amont of variation to the variation among geographic regions, the STRUCTURE analysis was able to detect strong clustering of genotypes within many populations. Data also indicated that gene flow is present among populations and that genetic diversity is not variable among the three geographic regions: Northeast (New Hanover County North Carolina, Norfolk County Virgina), Southeast (Everglades Florida, Florida Keys, Sanibel Island Florida), and Southwest (Cameron Parish Louisiana, Brazoria County Texas, Cameron County Texas, Willacy County Texas, San Patricio County Texas, Galveston County Texas, and Tamaulipas Mexico). Additionally most populations harbored unique alleles, indicating populations are genetically unique. A surprising exception was the lack of unique alleles in the Florida Keys population. This population is geographically isolated from mainland marsh rice rats and has developed distinguishing physical and behavioral traits that originally caused Spitzer and Lazell (1978) to classify it as a separate species. These factors indicate a high likelihood of genetic uniqueness. Also Wang et al. (2005) found four unique alleles across six loci in 18 individuals of the Keys rice rat population compared to 55 individuals from the Everglades. Perhaps the smaller sample size in this study did not allow for any unique alleles to be detected. Two of the seven populations, distributed along the Gulf Coast had unique genetic diversity. The Brazoria County and Galveston

County Texas populations may be larger than the others, enabling unique alleles to persist over many generations.

Populations from Willacy County and Cameron County Texas did not have any unique alleles and genetic diversity was relatively low in these two populations. Though data suggests moderate gene flow between these two populations, gene flow may be very low between these populations and other populations as indicated by higher F_{ST} values (Table 4.7). This lack of new alleles could be causing low genetic diversity within these two populations. An environmental influence, as well as demographic, may be maintaining diversity at a relatively low level. Perhaps an environmental factor favors genotypes adapted to the local habitat, while immigrants with foreign genotypes do not survive well in the local environment. Geography may also be affecting these populations along with the Tamaulipas Mexico population. These three populations are all south of the Texas Coastal Sand Plain. Though some marshes exist on the coast, and semi-perminant wetlands are found scattered throughout the grasslands of the Coastal Sand Plain (Fulbright et al. 1990), dispersal may be limited through this geologic feature. Much of the grassland is used for grazing livestock, disturbing the natural ecosystem (Diamond and Fulbright 1990). This anthropogenic impact may also play a role in isolating these populations south of the Texas Coastal Sand Plain.

Populations from the Everglades Florida, Sanibel Island Florida, and Tamaulipas Mexico were in HWE. Significant deviations from HWE uncovered in the other populations are all due to lower observed heterozygosity than expected, indicating high levels of homozygosity. This may be due to inbreeding within these populations as indicated by the significant inbreeding coefficients. However, deviations from HWE are

not consistent across all loci, as would be expected if inbreeding caused these deviations (Selkoe and Toonen 2006). Low heterozygosity can also be caused by a Wahlund effect, an indication of structuring within populations, by genetic drift, or by selection. But these too should affect all loci (Selkoe and Toonen 2006). The deviations from HWE detected in these marsh rice rat populations may not be biologically real and are likely caused by undetected null alleles or large allele drop out. Screening of more individuals from these populations would increase the statistical power of HWE calculations and could show that these populations are actually in HWE.

Many of the marsh rice rat populations screened in this study showed LD. Linkage disequilibrium may be found in low levels in natural populations (Hartl and Clark 1997). Therefore, the linkage disequilibrium detected in the populations of this study is likely real, rather than an effect of sampling or analysis. LD can be caused by genetic drift in small populations, selection, by the rejoining of previously differentiated populations, and can arise when a populations is re-expanding after a population bottleneck (Hansson 2010). A genetic bottleneck was detected in only two populations, the Cameron Parish Louisiana population and the Brazoria County population, which may explain the LD among loci in these populations.

Varying levels of gene flow among marsh rice rat populations were detected in this study. Gene flow was estimated to be very high between some eastern and western population pairs, however this is unlikely to be real. The physical presence of dispersal between eastern and western populations has not been verified, so indications of gene flow detected within this study suggest a genetic signature caused by homoplasy. Eastern and western marsh rice rat populations have been separated for a relatively long time, as

long as two million years (Chapter Three). This long separation has allowed for back mutations within microsatellite loci to arise causing the true amount of genetic divergence between the two groups to be masked. No conclusions can be made regarding the present genetic connectivity between these two geographic groups because both mitochondrial and nuclear DNA sequences show a distinct separation between them (Hanson et al. 2010). If only mtDNA showed this divergence, gene flow could be inferred between eastern and western populations using the microsatellite loci. Genetic patterns uncovered by nuclear markers can differ from patterns uncovered from mtDNA (Yang and Kenagy 2009). However, nuclear sequence data has shown great differentiation between eastern and western populations. Further study of the genetic suture zone between eastern and western marsh rice rat populations is necessary to determine if levels of gene flow detected by microsatellite markers are the result of homoplasy or recent gene flow not detected by the DNA sequence data.

Gene flow is happening among populations within the eastern and western groups. An isolation by distance effect may be causing the generally higher F_{ST} values estimated between the populations from the Northeast (North Carolina and Virginia) and the Southeast (Florida). The two island populations in Florida each had moderate levels of gene flow with the Everglades population (Sanibel Island $F_{ST} = 0.078$, Florida Keys $F_{ST} = 0.142$). Though microsatellites evolve relatively quickly the genetic connectivity detected between island and mainland populations may be due to a recent divergence rather than because gene flow is still present (Holsinger and Weir 2009). Current gene flow between the Everglades and Keys population is unlikely due to the geographic distance between them. Also, mtDNA sequence data showed divergence of these island

populations supporting that gene flow is restricted or not present at all (Chapter Two and Three). Studies of dispersal in these island populations are required to determine the physical presence of gene flow.

Low to high levels of gene flow were detected among western populations. As mentioned above, moderate levels of gene flow were detected between Willacy County Texas, Cameron County Texas, and Tamaulipas Mexico, though little to no gene flow was detected between the Willacy County Texas population and other western populations. As mentioned above, these three populations are all in geographic proximity to one another, south of the Texas Coastal Sand Plain. Moderate levels of gene flow were detected between Cameron County Texas and other western populations, as well as between the Tamaulipas Mexico population and other western populations. However, the Cameron County Texas and Tamaulipas Mexico populations could be recently isolated from other western populations. High gene flow was detected among the other western populations along the Gulf Coast indicating that wetlands are connected along this region, and potential isolation from anthropogenic habitat fragmentation is minimal.

Geographic structuring of genetic diversity was detected among most populations. The strong clustering among individuals from the Sanibel Island Florida population, the Lower Keys Florida population, Norfolk County Virginia population, Cameron County Texas population, Galveston County Texas population, Willacy County Texas population, and the Tamaulipas Mexico population indicates that these populations are genetically distinct. In the case of the two island populations this geographic structuring may be due to isolation from mainland marsh rice rat populations, providing further support that these populations are not experiencing current gene flow with

mainland populations. The Everglades Florida individuals and the New Hanover County North Carolina individuals may not have clustered together because high genetic variation exists within them. The Everglades population is likely much larger than other populations because this wetland is protected as a National Park (Richardson 2009). This habitat protection combined with the vast size of the wetland has left populations of small mammal species free from environmental limitations experienced elsewhere (Beckmann, personal communication). A large, unrestricted population is likely to have greater genetic variation than a smaller population. Alternatively, within the New Hanover County population, low structuring may be due to the inclusion of samples separated temporally, indicating that genetic variation may have changed over time.

The three western populations that clustered together correctly in the structure analysis (Tamaulipas Mexico, Cameron County and Willacy County Texas) are geographically close to one another (Figure 4.1). Though gene flow is occurring among these three populations, genetic variation is relatively low, and no unique alleles were detected in two of these populations, they are still genetically differentiable. The moderate F_{ST} and R_{ST} values detected among these populations could still be found if gene flow has been recently restricted. The mtDNA sequence analyses of Chapters Two and Three also showed genetic divergence among these populations and low genetic diversity within these populations. These three populations could be isolated from other populations by the topography of the Texas Coastal Sand Plain, though marshes do exist along the coast (Fulbright et al. 1990). Genetic differentiation could also be due to restricted gene flow among geographically close wetlands that have been disconnected due to human development. The identification of populations that are changing due to

anthropogenic factors is the first step to protecting marsh rice rats and their critical wetland habitat.

Conclusions

The results of this study suggest that the marsh rice rat's dispersal ability is affecting its genetic structure, but habitat disconnection caused by anthropogenic factors may be opposing the species' natural gene flow in some regions. Although levels of gene flow among eastern and western populations cannot be determined, there are moderate levels of gene flow among populations within each group. However, some populations remain genetically distinct due to geographic isolation, either by natural or anthropogenic causes. The sampling limitations within this study may be affecting the patterns uncovered from these data. Microsatellite studies typically sample more than twenty individuals from a population and most of the statistical tests carried out in this study are more powerful with a higher sample size. Working with natural populations is difficult; sampling is limited by the behavior and distribution of the organism being studied, as well as the finite time and effort for sampling available to the researcher. It should also be noted that samples are only a subset of the true population; a different subset from the same population could produce different results.

Despite the small sample size of this study, microsatellite analyses of marsh rice rat populations revealed similar genetic patterns to the mtDNA phylogeography study in Chapter Three. The genetic patterns and levels of diversity within this species are now better understood. Conservation biologists can use this data when creating management plans for wetlands in the southeastern United States and developing conservation plans

for other small mammal species. The patterns uncovered in this study may guide research of similar small mammal species. This study leads to a more complete understanding of intra-specific population processes within small mammals, which in turn will help evolutionary biologists begin to unravel some of the many complexities of speciation.

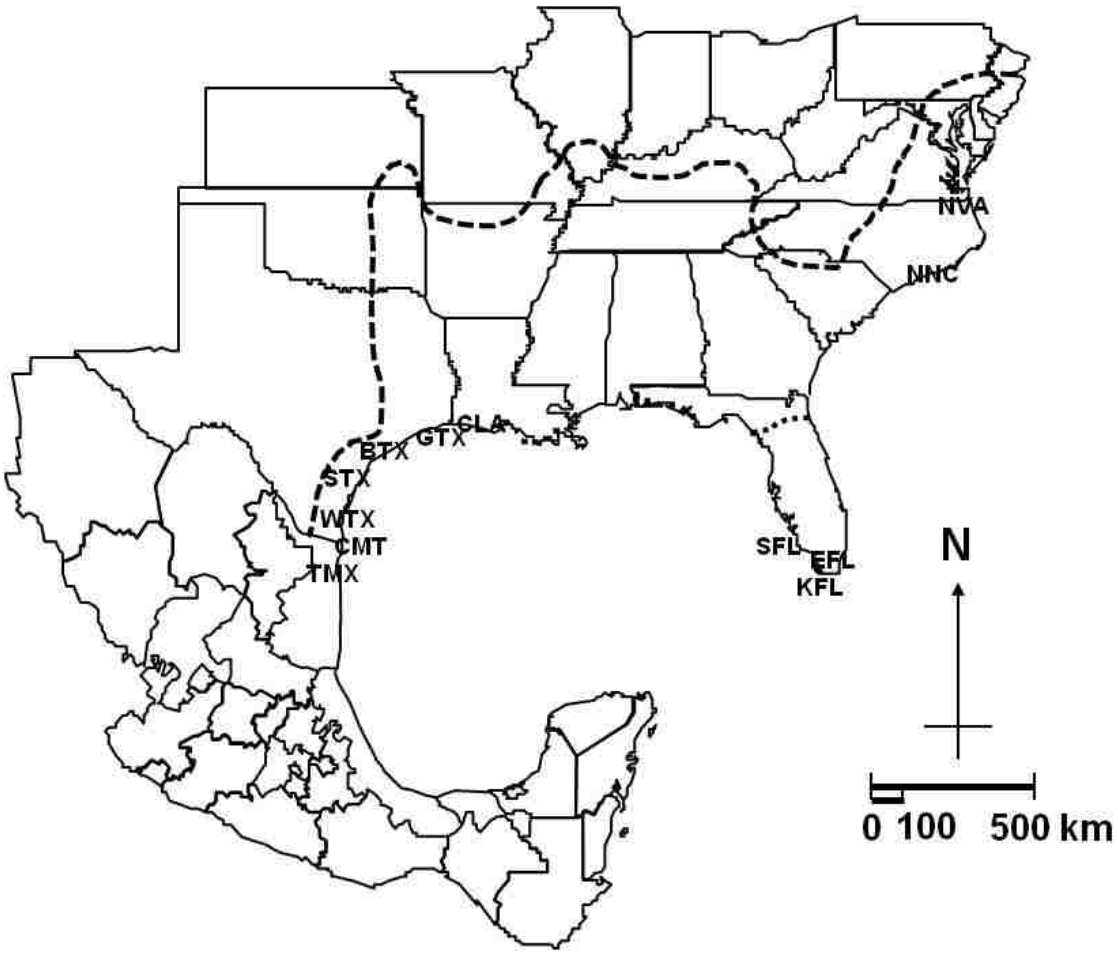


Figure 4.1. Distribution of the marsh rice rat and locations of the 12 populations used in this study. Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), KFL (Lower Keys, Monroe County, Florida), SFL (Sanibel Island, Lee County, Florida), NNC (New Hanover County, North Carolina), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, Texas), GTX (Galveston County, Texas), and TMX (Matamoros, Tamaulipas, Mexico).

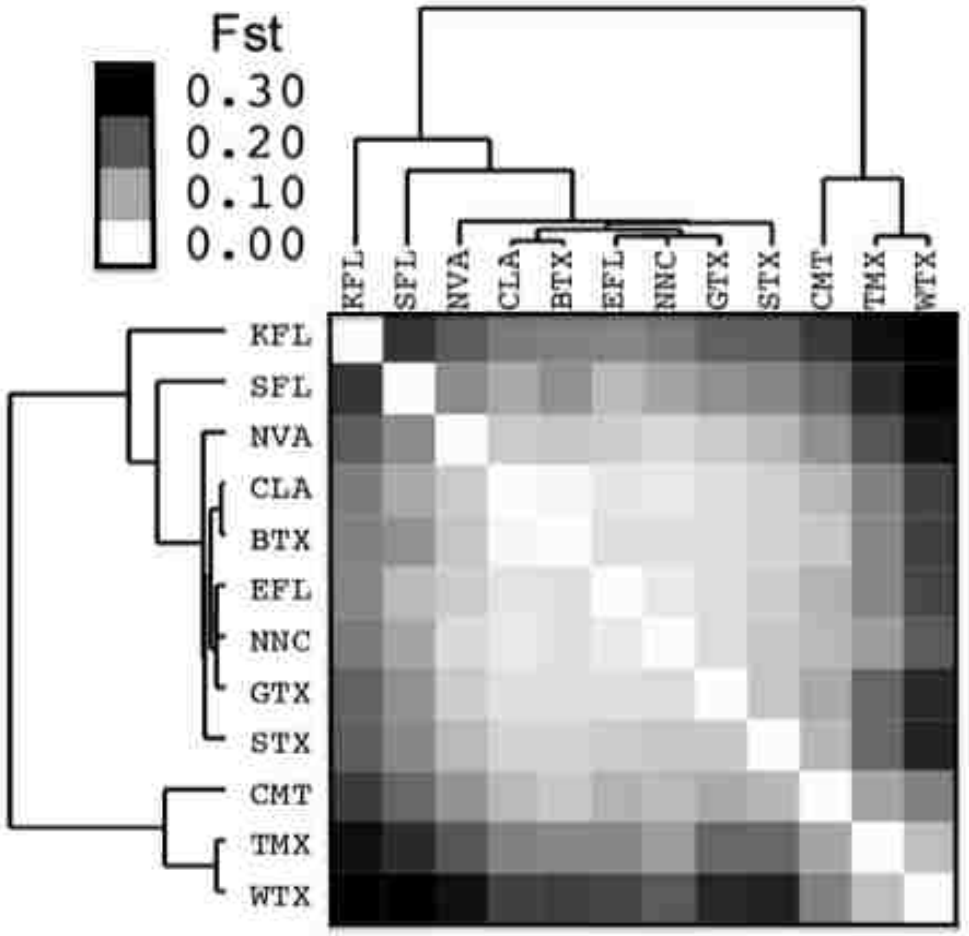


Figure 4.2. Hierarchical cluster analysis of F_{ST} values among populations. Populations with the most similar F_{ST} values were clustered together. F_{ST} values indicate the level of gene flow among populations. Values between 0 and 0.05 indicate high gene flow, between 0.05 and 0.15 moderate gene flow, and between 0.15 and 0.25 low gene flow. Values above 0.25 indicate the absence of gene flow. Some eastern and western populations appear to have more gene flow than they actually may have due to homoplasy. Some populations have identical alleles due to chance mutations, instead of gene flow or common ancestry because of the fast rate of evolution in microsatellite loci. Population abbreviations are the same as in Figure 4.1.

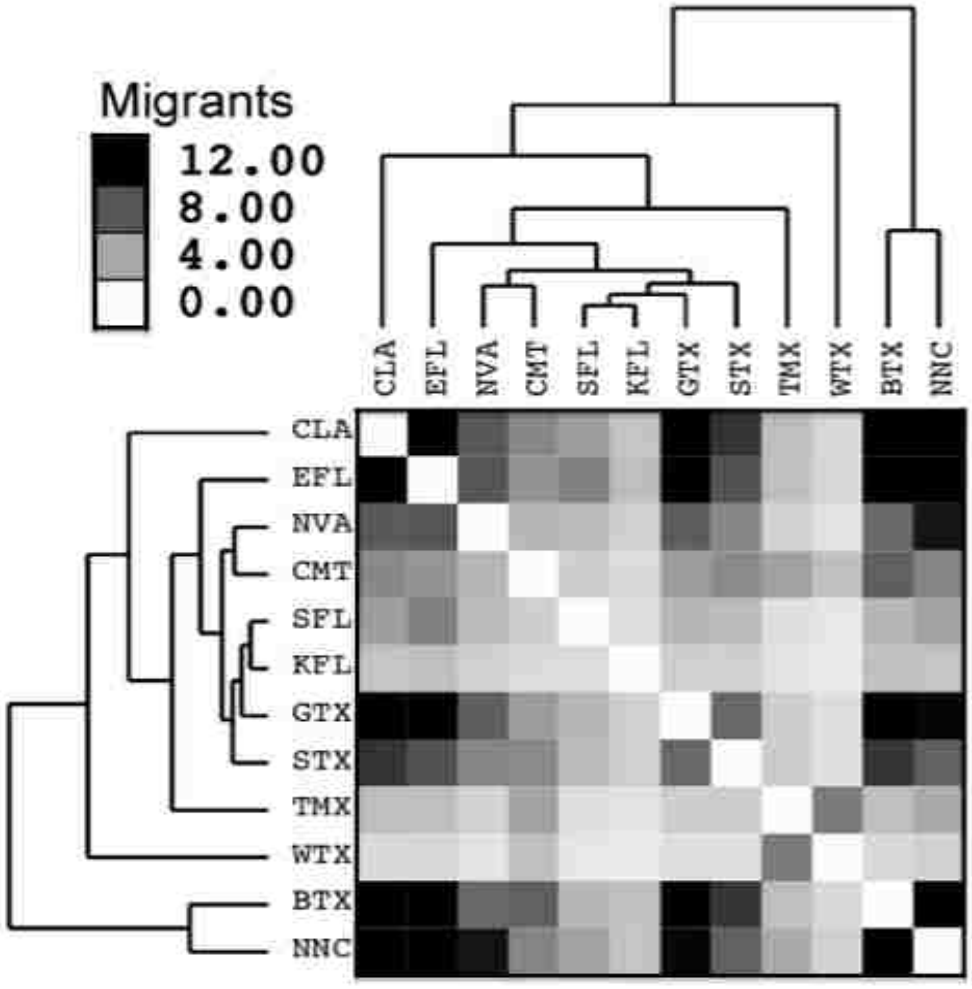


Figure 4.3. Hierarchical cluster analysis of the estimated number of migrants (M) among populations. Populations with the most similar M values were clustered together. Some eastern populations and some western populations appear to have more migrants than they really may have due to homoplasy. Because microsatellite loci evolve quickly, populations may have identical alleles due to chance mutation, instead of gene flow or common ancestery. Population abbreviations are the same as in Figure 4.1.

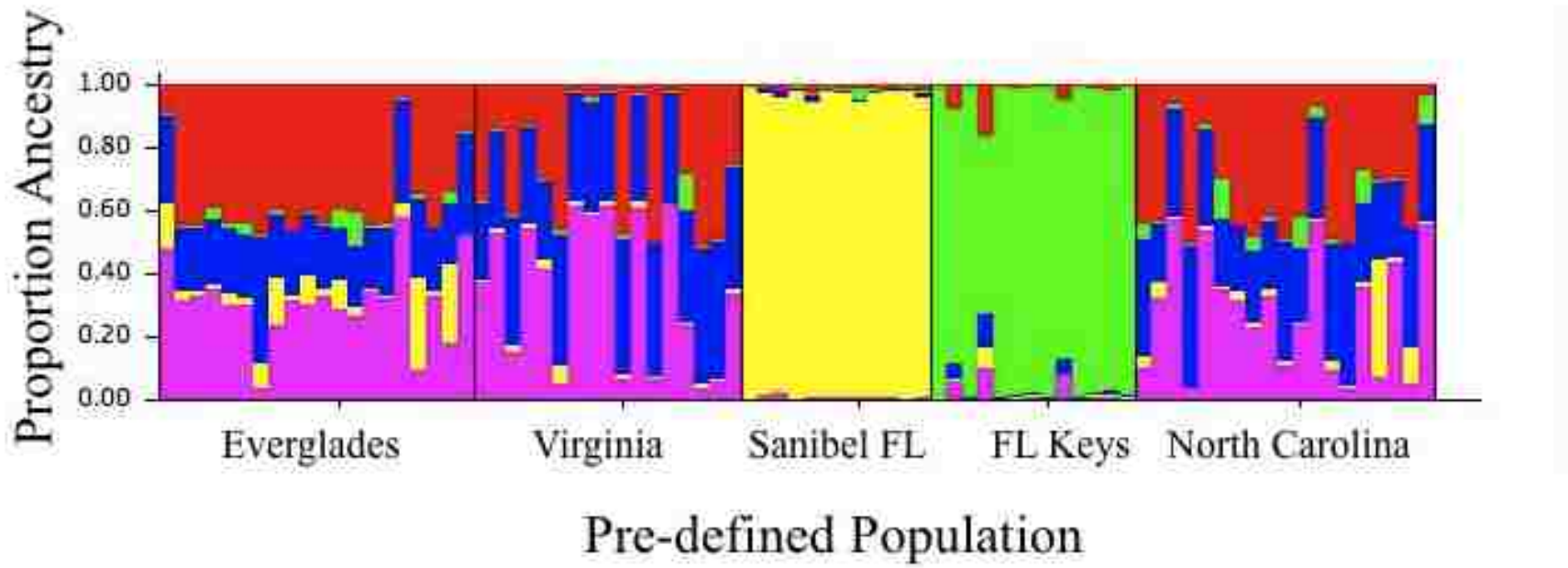


Figure 4.4. STRUCTURE analysis of the five eastern populations (Pritchard et al 2000, Pritchard and Wen 2003). Individuals from the Sanibel Island Florida population and the Florida Keys population clustered together strongly. Though the analysis correctly estimated the number of populations from which individuals came from (K = 5), individuals from the Everglades Florida population, the Virginia population, and North Carolina population did not cluster together. This indicates mixed ancestry of these individuals and the presence of gene flow among these three populations. This analysis supports the presence of little to no gene flow between the Sanibel Island Florida population and other populations, as well as between the Florida Keys population and the other populations.

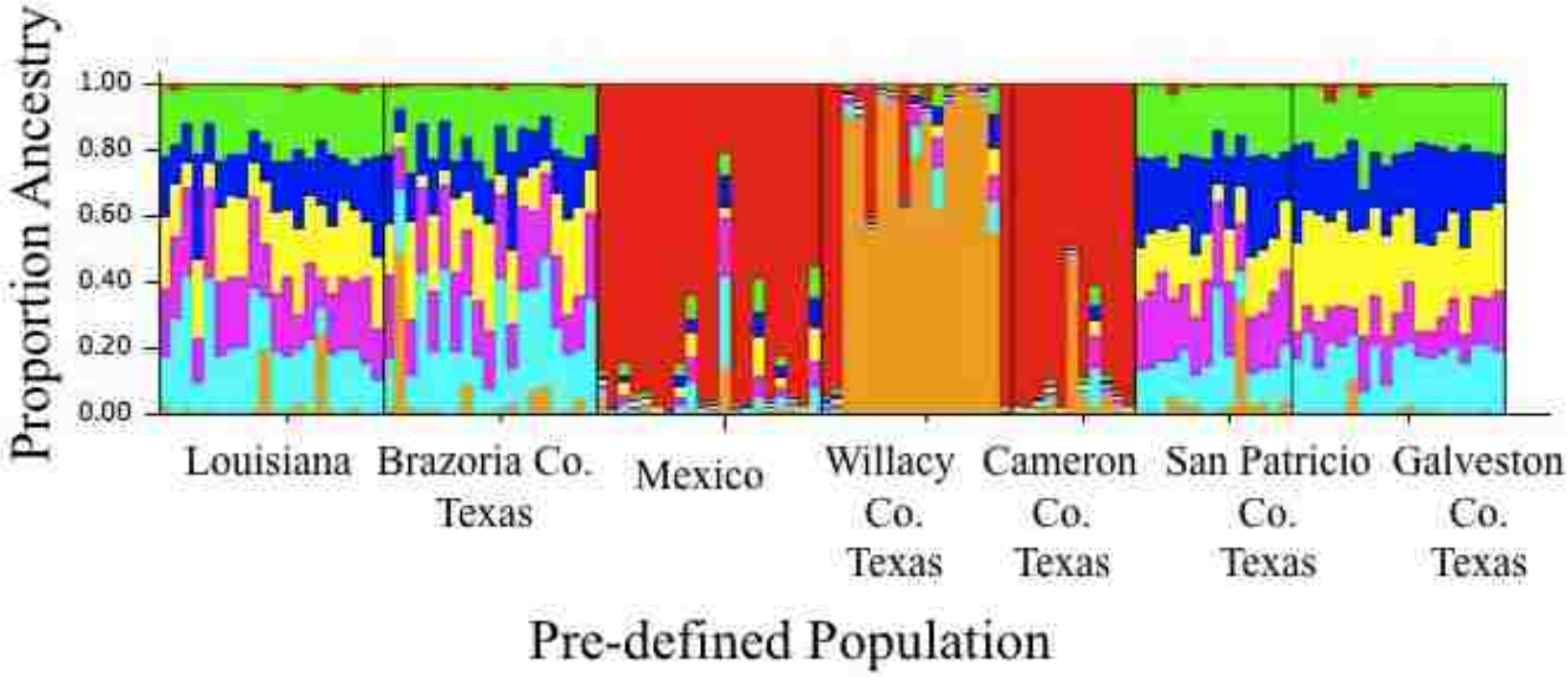


Figure 4.5. STRUCTURE analysis of the seven western populations (Pritchard et al 2000, Pritchard and Wen 2003). Individuals from the Willacy County Texas population, the Cameron County Texas population, and the Tamaulipas Mexico population clustered together strongly. Individuals from Mexico and Cameron County Texas share the same ancestry as indicated by the red color. Though the analysis correctly estimated the number of populations from which individuals came from (K = 7), individuals from Louisiana, Brazoria County Texas, San Patricio County Texas, and Galveston County Texas did not cluster together strongly. This indicates mixed ancestry of these individuals and the presence of gene flow among these four populations. This analysis supports the presence of little to no gene flow between Willacy County Texas and the other populations. Strong gene flow exists between the populations from Mexico and Cameron County Texas, but little gene flow is present between these two populations and other populations.

Table 4.1. Microsatellite primer concentrations and annealing temperatures used in this study. Annealing temperatures are as in Wang et al. 2000.

Locus	Dye-Label	Primer Concentration	Annealing Temperature
OryAAT03	NED yellow	0.3 μmol	55^{o}
OryAAT10	VIC green	0.18 μmol	53^{o}
OryAAT16	NED yellow	0.2 μmol	50^{o}
OryAAT21	VIC green	0.18 μmol	55^{o}
OryAAT26	PET red	0.2 μmol	50^{o}
OryAAT28	6-FAM blue	0.2 μmol	55^{o}
OryAAT40	PET red	0.2 μmol	55^{o}
OryAAT60	6-FAM blue	0.12 μmol	53^{o}
OryAAT64	6-FAM blue	0.4 μmol	50^{o}

Table 4.2. Number of alleles per locus for populations of *Oryzomys palustris* throughout the species' range. Within all 12 populations there were 144 alleles. Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), KFL (Lower Keys, Monroe County, Florida), SFL (Sanibel Island, Lee County, Florida), NNC (New Hanover County, North Carolina), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, Texas), GTX (Galveston County, Texas), and TMX (Matamoros, Tamaulipas, Mexico). Numbers of individuals in each population are in parentheses.

Locus	EFL (20)	KFL (13)	SFL (12)	NNC (19)	NVA (17)	CLA (20)	BTX (19)	CMT (17)	WTX (11)	STX (14)	GTX (19)	TMX (20)	Mean	Total Number
OryAAT03	10	7	3	9	6	8	10	9	2	6	7	7	7.0	16
OryAAT10	10	5	6	9	6	9	12	6	5	8	8	8	7.667	17
OryAAT16	9	3	6	8	7	8	7	4	2	6	9	5	6.167	12
OryAAT21	10	5	6	12	9	10	9	8	5	10	7	8	8.250	16
OryAAT26	13	5	6	8	8	7	8	7	6	7	9	5	7.417	17
OryAAT28	11	6	3	9	6	8	13	6	3	5	9	7	7.167	18
OryAAT40	7	5	6	8	6	9	10	6	6	10	11	4	7.333	13
OryAAT60	9	4	3	10	6	12	11	7	5	13	13	5	8.167	18
OryAAT64	9	5	5	5	7	9	10	6	2	6	7	5	6.333	17
Mean	9.778	5.0	4.889	8.667	6.778	8.889	10.0	6.556	4.0	7.889	8.889	6.0	7.278	16

Table 4.3. Observed and expected heterozygosity (H_O/H_E) for each locus within each population. Significant differences are distinguished with an asterisks ($p < 0.05$). Loci that were still not in Hardy-Weinberg equilibrium after adjusting allele frequencies for null alleles are in bold. Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), KFL (Lower Keys, Monroe County, Florida), SFL (Sanibel Island, Lee County, Florida), NNC (New Hanover County, North Carolina), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, Texas), GTX (Galveston County, Texas), and TMX (Matamoros, Tamaulipas, Mexico).

	AAT03	**AAT10**	**AAT16**	**AAT21**	**AAT26**	**AAT28**	**AAT40**	**AAT60**	**AAT64**	**Mean**
EFL	0.55/0.88*	0.90/0.87	0.90/0.89	0.85/0.89	0.85/0.89	0.70/0.87	0.75/0.85	0.90/0.82	0.50/0.81*	0.77/0.86
KFL	0.92/0.80	0.85/0.63	0.54/0.43	0.62/0.82	0.38/0.46	**0.54/0.75***	0.69/0.71	0.69/0.58	0.09/0.82*	0.59/0.66
SFL	0.67/0.69	0.75/0.75	0.67/0.65	0.83/0.80	0.50/0.83*	0.58/0.68	0.92/0.74	0.50/0.54	0.17/0.70*	0.62/0.71
NNC	0.32/0.87*	**0.53/0.87***	0.88/0.88	0.89/0.88	**0.46/0.86***	0.74/0.85	0.79/0.87	0.94/0.89	0.11/0.72*	0.63/0.86
NVA	0.65/0.76	0.65/0.70	0.8240.75	0.94/0.88	0.50/0.83*	0.82/0.73	**0.77/0.78***	0.94/0.79	0.38/0.82*	0.70/0.78
CLA	0.50/0.84*	0.70/0.86	0.70/0.84	0.85/0.90	**0.65/0.80***	0.95/0.82	0.85/0.88	0.94/0.79	0.47/0.87*	0.72/0.86
BTX	0.42/0.87*	0.68/0.90	0.84/0.82	1/0.89	0.42/0.78*	0.84/0.91	1/0.87	0.95/0.90	**0.71/0.88***	0.76/0.87
CMT	0.53/0.88*	0.41/0.75*	0.76/0.75	0.76/0.82	**0.59/0.70***	0.35/0.79*	**0.59/0.78***	0.82/0.82	0.47/0.76*	0.59/0.78
WTX	**0/0.17***	0.36/0.71*	0.27/0.45	0.82/0.68	0.8/0.78	**0.09/0.26***	1/0.72	0.64/0.71	0.09/0.25	0.45/0.53
STX	0.71/0.83	0.50/0.83*	0.50/0.74	1/0.87	0.71/0.83	0.71/0.78	0.86/0.90	0.93/0.95	**0.50/0.74***	0.71/0.83
GTX	0.47/0.75*	0.89/0.80	0.95/0.89	0.95/0.85	**0.74/0.82***	0.95/0.80	0.89/0.90	0.84/0.88	0.37/0.79*	0.78/0.83
TMX	0.70/0.77	0.32/0.73*	0.55/0.63	0.80/0.80	0.67/0.76	0.3/0.62*	0.6/0.65	0.45/0.60	0.35/0.57*	0.53/0.68
Mean H_E	0.76	0.78	0.73	0.84	0.78	0.74	0.80	0.78	0.73	0.77

Table 4.4. Frequencies of null alleles at each locus within each *Oryzomys palustris* population. Loci AAT16, AAT21, AAT40, and AAT60 did not any have null alleles. Allele frequencies were estimated using Brookfield's (1996) method. Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), SFL (Sanibel Island, Lee County, Florida), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), KFL (Lower Keys, Monroe County, Florida), TMX (Matamoros, Tamaulipas, Mexico), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, NNC (New Hanover County, North Carolina), Texas), and GTX (Galveston County, Texas).

Population	AAT03	AAT10	AAT26	AAT28	AAT64
EFL	0.1655	-	-	-	0.1597
NVA	-	-	0.2488	-	0.3239
SFL	-	-	0.166	-	0.3
CLA	0.177	-	-	-	0.2816
BTX	0.2303	0.1019	0.1921	-	-
KFL	-	-	-	-	0.5531
TMX	-	0.3132	-	0.1919	0.1332
CMT	0.1754	0.184	-	0.2333	0.3416
WTX	-	0.1892	-	-	-
STX	-	0.1648	-	-	-
NNC	0.2889	0.4466	0.5731	-	0.4169
GTX	0.1495	-	-	-	0.2281

Table 4.5. Loci remaining in linkage disequilibrium (LD) after using a sequential Bonferroni correction for multiple comparisons. There were none in the Willacy County Texas population. Adjusted P = 0.0014 for 35 comparisons.

Population	Loci pairs in LD
Everglades Florida	AAT03 and AAT21 (P = 0)
Norfolk County Virginia	AAT03 and AAT28 (P = 0)
Sanibel Island Florida	AAT03 and AAT28 (P = 0)
Cameron Parish Louisiana	AAT03 and AAT28, AAT03 and AAT60 (P = 0, P = 0.001)
Brazoria County Texas	AAT03 and AAT28 (P = 0)
Lower Keys Florida	AAT03 and AAT28 (P = 0)
Tamaulipas Mexico	AAT03 and AAT10, AAT03 and AAT28, AAT10 and AAT60 (P < 0.001, P = 0, P = 0, P = 0)
Cameron County Texas	AAT03 and AAT28, AAT10 and AAT60, AAT21 and AAT26 (P < 0.001, P = 0, P < 0.001)
San Patricio County Texas	AAT03 and AAT28, AAT10 and AAT60 (P = 0, P < 0.001)
New Hanover County North Carolina	AAT10 and AAT16, AAT03 and AAT28, AAT40 and AAT60, AAT03 and AAT64 (P = 0.001, P = 0, P = 0, P < 0.001)
Galveston County Texas	AAT03 and AAT28 (P = 0)

Table 4.6. Genetic diversity estimates within 12 populations of *Oryzomys palustris*. Measures are average gene diversity (π), Nei's gene diversity (H_E), mean allelic richness, and population specific F_{IS}, which is an estimate of inbreeding. Significant values are marked with an asterisk ($P < 0.05$). Populations are Norfolk County Virginia (NVA), New Hanover County North Carolina (NNC), Lower Keys Florida (KFL), Cameron Parish Louisiana (CLA), Brazoria County Texas (BTX), Everglades, Florida (EFL), Willacy County Texas (WTX), Galveston County Texas (GTX), Cameron County Texas (CMT), Tamaulipas Mexico (TMX), San Patricio County Texas (STX), Sanibel Island Florida (SFL).

Population	Nei's Average Gene Diversity (H_E)	Mean Allelic Richness (R)	F_{IS}
NVA	0.786	6.037	0.112*
NNC	0.863	7.629	0.096*
KFL	0.669	4.713	0.130*
CLA	0.865	7.739	0.152*
BTX	0.870	8.280	0.113*
EFL	0.862	7.911	0.080
WTX	0.529	3.908	0.218*
GTX	0.832	7.363	0.245*
CMT	0.790	5.738	0.134
TMX	0.685	4.775	0.140*
STX	0.831	7.174	0.199*
SFL	0.720	4.926	0.058

Table 4.7. Pairwise F_{ST} values among 12 populations of *Oryzomys palustris* from nine microsatellite loci. Significant differences are distinguished with an asterisk ($P < 0.05$). Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), SFL (Sanibel Island, Lee County, Florida), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), KFL (Lower Keys, Monroe County, Florida), TMX (Matamoros, Tamaulipas, Mexico), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, NNC (New Hanover County, North Carolina), Texas), and GTX (Galveston County, Texas).

	EFL	**NVA**	**SFL**	**CLA**	**BTX**	**KFL**	**TMX**	**CMT**	**WTX**	**STX**	**NNC**	**GTX**
EFL	0											
NVA	0.059*	0										
SFL	0.078*	0.135*	0									
CLA	0.028*	0.060*	0.100*	0								
BTX	0.035*	0.066*	0.130*	0.010	0							
KFL	0.142*	0.191*	0.238*	0.153*	0.146*	0						
TMX	0.140*	0.198*	0.250*	0.146*	0.141*	0.284*	0					
CMT	0.089*	0.128*	0.180*	0.083*	0.063*	0.230*	0.104*	0				
WTX	0.215*	0.280*	0.317*	0.222*	0.225*	0.375*	0.075*	0.147*	0			
STX	0.058*	0.082*	0.140*	0.050*	0.050*	0.191*	0.177*	0.084*	0.259*	0		
NNC	0.023*	0.043*	0.104*	0.024*	0.035*	0.155*	0.115*	0.081*	0.194*	0.063*	0	
GTX	0.037*	0.062*	0.130*	0.035*	0.038*	0.186*	0.179*	0.099*	0.254*	0.065*	0.040*	0

Table 4.8. Pairwise R_{ST} values among 12 populations of *Oryzomys palustris* from nine microsatellite loci. Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), SFL (Sanibel Island, Lee County, Florida), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), KFL (Lower Keys, Monroe County, Florida), TMX (Matamoros, Tamaulipas, Mexico), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, NNC (New Hanover County, North Carolina), Texas), and GTX (Galveston County, Texas).

	EFL	NVA	SFL	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
EFL	0											
NVA	0.063	0										
SFL	0.085	0.156	0									
CLA	0.029	0.064	0.111	0								
BTX	0.036	0.071	0.150	0.010	0							
KFL	0.166	0.236	0.312	0.181	0.171	0						
TMX	0.176	0.247	0.334	0.171	0.165	0.396	0					
CMT	0.098	0.147	0.219	0.091	0.067	0.298	0.116	0				
WTX	0.273	0.388	0.465	0.285	0.290	0.601	0.081	0.173	0			
STX	0.061	0.089	0.163	0.052	0.053	0.236	0.215	0.091	0.35	0		
NNC	0.024	0.045	0.116	0.024	0.037	0.183	0.130	0.088	0.241	0.067	0	
GTX	0.039	0.066	0.150	0.036	0.040	0.228	0.218	0.110	0.340	0.070	0.043	0

Table 4.9. The absolute number of migrants (M) among 12 populations of *Oryzomys palustris* estimated from nine microsatellite loci (M = 2nm). M is estimated from F_{ST} assuming migration-drift equilibrium (Slatkin 1991). Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), SFL (Sanibel Island, Lee County, Florida), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), KFL (Lower Keys, Monroe County, Florida), TMX (Matamoros, Tamaulipas, Mexico), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, NNC (New Hanover County, North Carolina), Texas), and GTX (Galveston County, Texas).

	EFL	**NVA**	**SFL**	**CLA**	**BTX**	**KFL**	**TMX**	**CMT**	**WTX**	**STX**	**NNC**	**GTX**
EFL												
NVA	7.949											
SFL	5.912	3.199										
CLA	17.306	7.783	4.503									
BTX	13.985	7.059	3.333	50.112								
KFL	3.019	2.116	1.600	2.760	2.924							
TMX	2.838	2.023	1.497	2.919	3.038	1.262						
CMT	5.010	3.391	2.283	5.513	7.446	1.675	4.300					
WTX	1.829	1.288	1.076	1.752	1.726	0.832	6.209	2.892				
STX	8.165	5.630	3.072	9.554	9.481	2.117	2.329	5.485	1.428			
NNC	21.006	11.007	4.296	20.450	13.608	2.734	3.856	5.683	2.076	7.434		
GTX	12.900	7.539	3.344	13.958	12.536	2.192	2.291	4.549	1.472	7.139	11.626	

Table 4.10. Analysis of molecular variance (AMOVA) for nine nuclear microsatellite loci from 12 *Oryzomys palustris* populations grouped by region. Regions are Northeast (Norfolk County Virginia, New Hanover County North Carolina), Southeast (Everglades and Sanibel Island Florida, Florida Keys), and Southwest (Cameron Parish Louisiana, Brazoria County, Cameron County, Willacy County, San Patricio County, and Galveston County, Texas, and Tamaulipas Mexico). Most of the genetic variation is attributable to within individuals and among individuals within populations.

Source of Variation	df	Sum of Squares	Variance Components	Percent of Variation
Among regions	2	46.255	0.045 (V_a)	1.15
Among populations Within regions	9	158.668	0.413 (V_b)	10.59
Among individuals		741.717	0.483 (V_c)	12.40
Within populations	189			
Within individuals	201	594.5	2.958 (V_d)	75.86
Total	401	1541.139	3.899	

Table 4.11. Proportion of individuals from each *Oryzomys palustris* population assigned to each of the nine clusters. Though samples came from 12 populations, STRUCTURE software (Pritchard et al. 2000) determined that individuals best fit into nine clusters. Each cluster number is given across the top of the table and N is the number of individuals in each given population. Values in bold represent the majority of individuals from a certain population that make up that cluster.

Given Population	1	2	3	4	5	6	7	8	9	N
EFL	0.312	0.146	0.092	0.112	0.028	0.013	0.082	0.191	0.024	20
NVA	0.056	0.058	0.107	0.020	0.010	0.011	**0.668**	0.056	0.014	17
SFL	0.010	0.009	0.010	**0.937**	0.005	0.004	0.010	0.009	0.005	12
CLA	**0.208**	**0.255**	0.122	0.037	0.035	0.009	0.079	**0.243**	0.012	20
BTX	0.184	**0.263**	0.130	0.013	0.072	0.008	0.056	**0.246**	0.028	19
KFL	0.019	0.016	0.011	0.009	0.022	0.005	0.019	0.017	**0.882**	13
TMX	0.025	0.021	0.007	0.024	0.019	**0.844**	0.032	0.022	0.007	20
CMT	0.012	0.015	0.016	0.007	**0.694**	0.220	0.009	0.015	0.012	17
WTX	0.014	0.016	0.007	0.007	0.064	**0.860**	0.011	0.017	0.005	11
STX	**0.228**	**0.307**	0.013	0.017	0.070	0.008	0.071	**0.280**	0.005	14
NNC	0.152	0.129	0.136	0.078	0.012	0.134	0.161	0.135	0.061	19
GTX	0.145	0.103	**0.520**	0.033	0.020	0.012	0.032	0.112	0.023	19

Chapter Five

General Conclusions

This dissertation thoroughly surveys the genetic architecture of the wetland dependent rodent, the marsh rice rat (*Oryzomys palustris*). The intra-specific systematics of the marsh rice rat disclosed that this species is split into two distinct eastern and western genetic groups. The great amount of genetic divergence between them calls for the western group to be elevated from a subspecies to a species, which would be named appropriately as *O. texensis*. Distinguishing cranial characters may exist between the two species, as indicated by past morphological analyses (Humphrey and Setzer 1989). However, morphology seems to be a poor indication of genetic variation within *Oryzomys* species. An intensive morphological analysis of *O. palustris* and *O. texensis* is needed to determine if diagnostic morphological characters exist. Study of ecological and behavioral differences between the two groups also will aid in establishing separate species status for eastern and western marsh rice rats. These data present the need for further study along the genetic suture zone to determine if these two groups are hybridizing and what, if any, present day factors are maintaining this distinct division. Dispersal and gene flow levels need to be studied with both an ecological and population genetics approach. This system makes for an intriguing study concerning speciation and hybridization.

Though six subspecies have been identified based on morphology, genetic data did not uncover six separate evolutionary lineages within the marsh rice rat. Based on the mtDNA genetic data, as well as morphology and natural history, I propose that only three subspecies are present within the marsh rice rat, *O. p. palustris*, *O. p. argentatus*, and *O.*

p. sanibeli. The first subspecies is comprised of mainland marsh rice rat populations west to the Alabama – Mississippi border, while the other two are distinct island populations in Florida. These two island subspecies may be isolated from mainland populations, further supporting their distinction as separate subspecies. Both mtDNA and nuclear microsatellite data support these two subspecies genetic distinction. Many islands exist along the Eastern Seaboard and Gulf Coast of the United States, so the existence of other island subspecies is not unrealistic and their presence should be determined.

The phylogeography of the marsh rice rat has been shaped by the past climatic and geoglogic history of the southeastern United States, as well as this species' wetland habitat specialization and over-water dispersal ability. The marsh rice rat's phylogeographic patterns are different from those of other small mammal species studied in the southeastern United States. For example, both the eastern woodrat (*Neotoma floridana*) and the white-tailed deer (*Odocoileus virginianus*) are geographically structured into three distinct groups of mtDNA haplotypes, instead of the two uncovered in the marsh rice rat (Ellsworth et al. 1994, Hayes and Harrison 1992). The Apalachicola River separates these species' eastern and western mtDNA clades, whereas the separation between the eastern and western clades of the marsh rice rat is further west along the border of Alabama and Mississippi. Like the marsh rice rat, eastern and western populations of the cotton rat (*Sigmodon hispidus*) are genetically divergent (Phillips et al. 2007). These two species are commonly found in the same habitat (Cameron and Kruchek 2005). However, the contact zone between eastern and western populations of the cotton rat is located in eastern Texas, farther west than the divide between eastern and western marsh rice rats.

Though a genetic discontinuity at the Apalachicola River exists for many studied species, there are a few which do not show this pattern. The ornate chorus frog (*Pseudacris ornate*) is a habitat specialist. Like the marsh rice rat this amphibian is dependant on wetlands, but unlike the marsh rice rat is a poor disperser (Beebee 2005). The ornate chorus frog did not exhibit a genetic discontinuity at the Apalachicola River; one of the three mtDNA clades found within this southeastern United States species, actually spanned the river (Degner et al. 2010). Four species in the southeastern United States show a phylogeographic pattern similar to the marsh rice rat, with their east-west genetic discontinuity not at the Apalachicola River. The sunfish (*Lepomis gulosus*; Bermingham and Avise 1986), water snakes (*Nerodia rhombifera* and *N. taxispilota*; Lawson 1987) and the Carolina chickadee (*Parus caroliniensis*; Gill et al. 1993) exhibit a genetic discontinuity at the Tombigbee River in western Alabama (Soltis et al. 2006). This discontinuity has been attributed to a Pliocene vicariance event, similar to the one that created the Apalachicola River discontinuity during the Pleistocene (Soltis et al. 2006).

Though inhabiting the same habitat, species can respond differently to climatic and geoglogic events depending on many biological factors, such as behavior and reproductive system (Stewart et al. 2010). The distribution and genetic diversity of the marsh rice rat was influenced by the Pleistocene glacial and interglacial cycles to some degree, however the major genetic patterns may have been formed earlier towards the end of the Pliocene and beginning of the Pleistocene, when climate was changing more quickly than this species may have persisted through previously. During the Pleistocene, local populations may have become extinct in areas where habitat became unsuitable, and

populations may have expanded into new habitat that became available with changing sea level. The extinction of marsh rice rat populations could have decreased the species overall genetic diversity, while newly formed populations may have increased it. Because of these fluctuations, present day genetic patterns within the marsh rice rat may only reflect the influence of the very last climate change of the Pleistocene. Biogeographic inferences such as these cannot be confirmed with complete certainty because supporting inferences are indirect. The evolutionary history of any organism is difficult to infer, but general hypotheses can be supported by both geologic and genetic data.

Genetic similarity and connectivity between eastern and western marsh rice rat populations detected by microsatellite markers is biased by chance mutations in these genetic regions that caused the same alleles to arise in different populations. This genetic similarity is due to homoplasy, not because of recent shared ancestry. Though marsh rice rats potentially have a great dispersal ability, most of the populations studied retained some level of genetic distinctness based on both mtDNA and nuclear microsatellite data, even those within the same geographic region. Restricted gene flow among populations could possibly create this pattern. Gene flow may be restricted by lack of connectivity among wetlands caused by human development.

Because of its dependency on wetland habitat, the marsh rice rat may be more vulnerable to climate change and anthropogenic effects than other rodent species. Populations may already be threatened due to habitat loss. Knowledge of the genetic relationships among populations of the marsh rice rat will be crucial in developing a

management plan for this species. Conservation will be more feasible and effective if the underlying genetic architecture of the species is known.

The effect of anthropogenic factors on genetic diversity has become a forefront issue in conservation. One in four mammal species is threatened with extinction (Schipper et al. 2008). Habitat loss is one of the main threats to the persistence of many mammal species. Habitat is shrinking worldwide, at an average rate of 1% per year, due to climate change and anthropogenic enchroachment (Balmford et al. 2003). Preserving habitat for species will aid in their conservation. However, preserving the evolutionary processes at work within a species ultimately may be more crucial, and more effective, in ensuring their future persistence (Mace and Purvis 2008). Phylogeographic studies will increase our understanding of the historical evolutionary forces, population dynamics, and climatic influences at work within species. Phylogeography may help to predict how species will respond to climate changes by testing hypotheses about how they reacted in the past. Studying a species' present genetic diversity and structure allows for unique populations to be identified. These populations can then be incorporated into a management plan that will help preserve genetic diversity within a species. Without these genetic data, a threatened or endangered species is likely to become inbred and lose a significant amount of its genetic diversity. With reduced genetic variation a species is less able to adapt to a changing environment and will become more susceptible to disease and extinction.

The current genetic structure and diversity within the marsh rice rat is influenced by local environmental variables, past geologic and climatic events, and more recently by anthropogenic habitat changes. The phylogeographic inferences of this study give a

glimpse into how climate change may affect this species. Genetic diversity and its distribution are forefront issues in conservation biology. A species will persist into the future only if diversity remains. Conservation biologists can use the genetic data generated by this phylogeographic study to create effective management plans for the marsh rice rat. Clearly island populations, eastern populations, and western populations need to be managed separately. The future of the marsh rice rat, as well as that of other wetland dependent species of the southeastern United States, will ultimately be determined by the fate of wetland habitat in the presence of a changing climate.

References

Allen, G. M. 1942. Extinct and vanishing mammals of the Western Hemisphere with the marine species of all oceans. American Committee for International Wildlife Protection, Special Publications 11**:**1 - 620.

Allen, J. A. 1894. On the mammals of Aransas County, Texas, with descriptions of new forms of *Lepus* and *Oryzomys*. Bulletin of the American Museum of Natural History 6**:**165 - 198.

Almeida, F. C., L. S. Maroja, M. A. M. Moreira, H. N. Seuánez and R. Cerqueira. 2005. Population structure and genetic variability of mainland and insular populations of the Neotropical water rat, *Nectomys squamipes* (Rodentia, Sigmodontinae). Genetics and Molecular Biology 28:693 - 699.

Alston, E. R. 1876. On two new species of *Hesperomys*. Proceedings of the Zoological Society of London 1876:775 - 757.

Álvarez-Castañeda, S. T. 2010. Phylogenetic structure of the *Thomomys bottae-umbrinus* complex in North America. Molecular Phylogenetics and Evolution 54:671 - 679.

Arbogast, B. S., R. A. Browne and P. D. Weigl. 2001.Evolutionary genetics and Pleistocene biogeography of North American tree squirrels (*Tamiasciurus*). Journal of Mammalogy 82:302 - 319.

Aris-Brosou, S. and L. Excoffier. 1996. The impact of population expansion and mutation rate heterogeneity on DNA sequence polmorphism. Molecular Biology and Evolution 13:494 - 504.

Austin, J. D., S. C. Lougheed and P. T. Boag. 2004. Discordant temporal and geographic patterns in the maternal lineages of eastern North American frogs, *Rana catesbeiana* (Ranidae) and *Pseudacris crucifer* (Hylidae). Molecular Phylogenetics and Evolution 32:799 - 816.

Avise, J. C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos 63:62 - 76.

Avise, J. C. 1996. Toward a regional conservation genetics perspective: phylogeography of faunas in the southeastern United States. *In* J. C. Avise and J. L. Hamrick (eds), Conservation Genetics: Case Histories from Nature. Chapman and Hall, New York, pp. 431 - 470.

Avise, J. C. 2000. Phylogeography: the History and Formation of Species. Harvard University Press, Cambridge, MA.

Avise, J. C. 2004. Molecular Markers, Natural History, and Evolution. Chapman & Hall, New York.

Avise, J. C. 2009. Phylogeography: retrospect and prospect. Journal of Biogeography 36:3 - 15.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb and N. C. Saunders. 1987. Intraspecific phylogeography - the mitochondrial-DNA bridge between population genetics and systematics. Annual Review of Ecology and Systematics 18:489-522.

Avise, J. C., C. Giblin-Davidson, J. Laerm, J. C. Patton and R. A. Lansman. 1979a. Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, *Geomys pinetis*. Proceedings of the National Academy of Sciences of the United States of America 76:6694 - 6698.

Avise, J. C., R. A. Lansman and R. O. Shade. 1979b. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. I. Population structure and evolution of the genus *Peromyscus*. Genetics 92:279 - 295.

Avise, J. C. and W. S. Nelson. 1989.Molecular genetic relationships of the extinct dusky seaside sparrow. Science 243:646 - 648.

Baird, S. F. 1857. Mammals: general report upon the zoology of the several Pacific railroad routes. Vol. 8, part 1. Pp 1 - 757 in Reports of explorations and surveys to ascertain the most practicable and economical route for a railroad from the Mississippi River to the Pacific Ocean. Washington, D. C., Senate Executive Document 78.

Baker, R. J. and R. D. Bradley. 2006. Speciation in mammals and the genetic species concept. Journal of Mammalogy 87:643 - 662.

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

Balloux, F. and N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. Molecular Ecology 11:155 - 165.

Balmford, A., R. E. Green and M. Jenkins. 2003. Measuring the changing state of nature. Trends in Ecology and Evolution 18:326 - 330.

Bangs, O. 1898. The land mammals of peninsular Florida and the coast region of Georgia. Proceedings of the Boston Society of Natural History 28:157 - 235.

Barbour, D. B. and S. R. Humphrey. 1982. Status of the silver rice rat (*Oryzomys argentatus*). Florida Scientist 45:112 - 116.

Bartlein, P. J., K. H. Anderson, P. M. Anderson, M. E. Edwards, et al. 1998. Paleoclimate simulations for North American over the past 21,000 years: features of the simulated climate and comaprisons with paleoenvironmental data. Quaternary Science Review 17:549 - 585.

Baum, D. 1992. Phylogenetic species concepts. Trends in Ecology and Evolution 7:1 - 2.

Beebee, T. J. C. 2005. Conservation genetics of amphibians. Heredity 95:423 - 427.

Beheregaray, L. B. 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Molecular Ecology 17:3754 - 3774.

Belfiore, N. M., L. Liu and C. Moritz. 2008. Multilocus phylogenetics of a rapid radiation in the Genus *Thomomys* (Rodentia: Geomyidae). Systematic Biology 57:294 - 310.

Bell, K. C., D. J. Hafner, P. Leitner and M. D. Matocq. 2010. Phylogeography of the ground squirrel subgenus *Xerospermophilus* and assembly of the Mojave Desert biota. Journal of Biogeography 37:363 - 378.

Bellinvia, E. 2004. A phylogenetic study of the genus *Apodemus* by sequencing the mitochondrial DNA control region. Journal of Zoological Systematics and Evolutionary Research 42:289 - 297.

Benson, D. L. and F. R. Gehlbach. 1979. Ecological and taxonomic notes on the rice rat (*Oryzomys couesi*) in Texas. Journal of Mammalogy 60:225 - 228.

Bermingham, E. and J. C. Avise. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. Genetics 113:939 - 965.

Bermingham, E. and C. Moritz. 1998. Comparative phylogeography: concepts and applications. Molecular Ecology 7:367 - 369.

Best, T. L. 1995. *Spermophilus mohavensis*. Mammalian Species 509:1 - 7.

Birky, C. W. 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annual Review of Genetics 35:125 - 148.

Blann, K. L., J. L. Anderson, G. R. Sands and B. Vondracek. 2009. Effects of agricultural drainage on aquatic ecosystems: A review. Critical Reviews in Environmental Science and Technology 39:909 - 1001.

Blaustein, R. J. 2008. Biodiversity Hotspot: The Florida Panhandle. Bioscience 58:784 - 790.

Bond, J. E., M. C. Hedin, M. G. Ramirez and B. D. Opell. 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trap-door spider *Aptostichus simus*. Molecular Ecology 10:899 - 910.

Bonvicino, C. R., P. R. Gonçalves, J. A. DeOliveira, L. F. B. DeOliveira and M. S. Mattevi. 2009. Divergence in *Zygodontomys* (Rodentia: Sigmodontinae) and distribution of Amazonian savannas. Journal of Heredity 100:322 - 328.

Bonvicino, C. R. and M. A. M. Moreira. 2001. Molecular phylogeny of the genus *Oryzomys* (Rodentia: Sigmodontinae) based on cytochrome b DNA sequences. Molecular Phylogenetics and Evolution 18:282 - 292.

Boone, J. L., M. H. Smith and J. Laerm. 1999. Allozyme variation in the cotton mouse (*Peromyscus gossypinus*). Journal of Mammalogy 80:833 - 844.

Borchsenius, F. 2009. FASTGAP 1.2. Department of Biological Sciences, University of Aarhus, Denmark. Published online at http://192.38.46.42/aubot/fb/FastGap_home.htm.

Bowler, D. E. and T. G. Benton. 2005. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. Biological Reviews 80:205 - 225.

Brant, S. V. and G. Ortí. 2003. Phylogeography of the northern short-taied shrew, *Blarina brevicauda* (Insectivora: Soricidae): Past fragmentation and postglacial recolonization. Molecular Ecology 12:1435 - 1449.

Briggs, G., T. Arkle and the Geological Society of America Southeastern Section. 1974. Carboniferous of the Southeastern United States: a symposium volume. Geological Society of America, Boulder, CO.

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

Brunjes, J. H. and W. D. Webster. 2003. Marsh rice rat, Oryzomys palustris, predation on Forster's Tern, *Sterna forsteri*, eggs in coastal North Carolina. Canadian Field-Naturalist 117:654 - 657.

Cameron, G. N. and B. L. Kruchek. 2005. Use of coastal wetlands by hispid cotton rats. The Southwestern Naturalist 50:397 - 402.

Carlsson, J. 2008. Effects of microsatellite null alleles on assignment testing. Journal of Heredity 99:616 - 623.

Carstens, B. C. and C. L. Richards. 2007. Integrating coalescent and ecological niche modleing in comparative phylogeography. Evolution 61:1439 - 1454.

Castleberry, S. B., T. L. King, P. B. Wood and M. Flord. 2002. Microsatellite DNA analysis of population structure in Allegheny Woodrats (*Neotoma magister*). Journal of Mammalogy 83:1058 - 1070.

Castro-Campillo, A., H. R. Roberts, D. J. Schmidly and R. D. Bradley. 1999. Systematic status of *Peromyscus boylii ambiguus* based on morphological and molecular data. Journal of Mammalogy 80:1214 - 1231.

Centeno-Cuadros, A., M. Delibes and J. A. Godoy. 2009. Phylogeography of southern water vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium? Molecular Ecology 18:3652 - 3667.

Chapman, F. M. 1893. Description of a new subspecies of *Oryzomys* from the Gulf states. Bulletin of the American Museum of Natural History 5:43 - 46.

Chapuis, M-P. and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. Molecular Biology and Evolution 24:621 - 631.

Church, S. A., J. M. Kraus, J. C. Mitchell, D. R. Church and D. R. Taylor. 2003. Evidence for multiple Pleistocene refugia in the postglacial expansion of the eastern tiger salamander, *Ambystoma tigrinum tigrinum*. Evolution 57:372 - 383.

Clegg, S. M., P. Hale and C. Moritz. 1998. Molecular population genetics of the red kangaroo (*Macropus rufus*): mtDNA variation. Molecular Ecology 7:679 - 686.

Clobert, J., E. Danchin, A. A. Dhondt and J. D. Nichols. 2001. Dispersal. Oxford University Press, Oxford, UK.

Comstock, K. E., N. Georgiadis, J. Pecon-Slattery, A. L. Roca, E. A. et al. 2002. Patterns of molecular genetic variation among Africa elephant populations. Molecular Ecology 11:2489 - 2498.

Cornuet, J. M. and G. Luikart. 1996. Description and power analysis of two tests for detecting population bottlenecks from allele frequency data. Genetics 144:2001 - 2014.

Cracraft, J. 1983. Species concepts and speciation analysis. Current Ornithology 1:159 - 187.

Crawford, J. C., Z. Liu, T. A. Nelson, C. K. Nielsen and C. K. Bloomquist. 2009. Genetic population structure within and between beaver (*Castor canadensis*) populations in Illinois. Journal of Mammalogy 90:373 - 379.

Croft, D. B. 1991. Home range of the red kangaroo *Macropus rufus*. Journal of Arid Environments 20:83 - 98.

Cronin, M. A. 1992. Intraspecific variation in mitochondrial DNA of North American cervids. Journal of Mammalogy 73:70 - 82.

Crouse, A. L. 2005. Genetic analysis of the endangered silver rice rat (*Oryzomys palustris natator*) and lower keys marsh rabbit (*Sylvilagus palustris hefneri*). M. S. thesis, Texas A&M University, 82 pp.

Dahl, T. E. 1990. Wetland losses in the United States: 1780's to 1980's. U.S. Fish and Wildlife Service, Washington, D.C.

Dalquest, W. W. 1962. The Good Creek formation, Pleistocene of Texas, and its fauna. Journal of Paleontology 36:568 - 582.

Dalquest, W. W. 1965. New Pleistocene formation and local fauna from Hardeman County, Texas. Journal of Paleontology 39:63 - 79.

Davidson, A. D., M. J. Hamilton, A. G. Boyer, J. H. Brown and G. Ceballos. 2009. Multiple ecological pathways to extinction in mammals. Proceedings of the National Academy of Sciences of the United States of America 106:10702 - 10705.

Davis, L. M., T. C. Glenn, D. C. Strickland, L. J. Guillette Jr., et al. 2002. Microsatellite DNA analyses support an east-west phylogeographic split of American Alligator populations. Journal of Experimental Zoology 294:352 - 372.

Degnan, J. H. and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology and Evolution 24:332 - 340.

Degner, J. F., D. M. Silva, T. D. Hether, J. M. Daza, and E. A. Hoffman. 2010. Fat frogs, mobile genes: unexpected phylogeographic patters for the ornate chorus frog (*Pseudacris ornata*). Molecular Ecology 19:2501 - 2515.

Degner, J. F., I. J. Stout, J. D. Roth and C. L. Parkinson. 2007. Population genetics and conservation of the threatened southeastern beach mouse (*Peromyscus polionotus niveiventris*): subspecies and evolutionary units. Conservation Genetics 8:1441 - 1452.

De Hoon, M. 2002. Cluster 3.0. Human Genome Center, University of Tokyo.

Delcourt, D. A. 1993. History, evolution, and organization of vegetation and human culture. *In* W. H. Martin, S. G. Boyce and A. C. Echternacht (eds), Biodiversity of the Southeastern United States: Lowland Terrestrial Communities. Wiley, New York.

Dewoody, J., J. D. Nason and V. D. Hipkins. 2006. Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes 6:951 - 957.

Diamond, D. D. and T. E. Fulbright. 1990. Contemporary plant communities of upland grassland of the Coastal San Plain, Texas. The Southwestern Naturalist 35:385 - 392.

Dieckmann, U., B. O'Hara and W. Weisser. 1999. The evolutionary ecology of dispersal. Trends in Ecology and Evolution 14:88 - 90.

Donovan, M. F., R. D. Semlitsch and E. J. Routman. 2000. Biogeography of the Southeastern United States: a comparison of salamander phylogeographic studies. Evolution 54:1449 - 1456.

Douglas, M. E., M. R. Douglas, G. W. Schuett, and L. W. Porras. 2009. Climate change and evolution of the New World pitviper genus *Agkistrodon* (Viperidae). Journal of Biogeography 36:1164 - 1180.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4:699 - 710.

Drummond, A. J., S. Y. W. Ho, N. Rawlence and A. Rambaut. 2007. A rough guide to BEAST 1.4.

Drummond, A. J. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214 - 222.

Ducroz, J-F., M. Stubbe, A. P. Saveljev, D. Heidecke, et al. 2005. Genetic variation and population structure of the Eurasian Beaver *Castor fiber* in Eastern Europe and Asia. Journal of Mammalogy 86:1059 - 1067.

Durka, W., W. Babik, J-F. Ducroz, D. Heidecke, et al. 2005. Mitochondrial phylogeography of the Eurasian beaver *Castor fiber* L. Molecular Ecology 14:3843 - 3856.

Dyke, A. S., J. T. Andrews, P. U. Clark, J. H. England, et al. 2002. The Laurentide and Innuitian ice sheets during the Last Glacial Maximum. Quaternary Science Review 21:9 - 31.

Eggert, L. S., C. A. Rasner and D. S. Woodruff. 2002. The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. Proceedings of the Royal Society of London B 269:1993 - 2006.

Ellsworth, D. L., R. L. Honeycutt, N. J. Silvy, J. W. Bickman and W. D. Klimstra. 1994. Historical biogeography and contemporary patterns of mitochondrial-DNA variation in white-tailed deer from the southeastern United States. Evolution 48:122 - 136.

El Mousadik, A. and R. J. Petit. 1996. High levels of genetic differentiation for allelic richness among populations of the argan tree (*Argania spinosa* (L). Skeels) endemic to Morocco. Theoretical and Applied Genetics 92:832 - 839.

Emslie, S. D. 1998. Avian community, climate, and sea-level changes in the Plio-Pleistocene of the Florida Peninsula. Ornithological Monographs 50:1 - 113.

Esher, J. L. Wolfe and Layne. 1978. Swimming behavior of rice rats (*Oryzomys palustris*) and cotton rats (*Sigmodon hispidus*). Journal of Mammalogy 59:551 - 558.

Estes-Zumpf, W. A., J. L. Rachlow, L. P. Waits and K. I. Warheit. 2010. Dispersal, gene flow, and population genetic structure in the pygmy rabbit (*Brachylagus idahoensis*). Journal of Mammalogy 91:208 - 219.

Evanno, G., S. Regnaut and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611 - 2620.

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

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

Fairbridge, R. W. 1974. The Holocene sea-level in South Florida. *In* P. J. Gleason (ed) The Holocene Sea-level in South Florida. Miami Geological Society, Memoir 2, pp. 223 - 232.

Faulkes, C. C., E. Verheyen, W. Verheyen, J. U. M. Jarvis and N. C. Bennett. 2004. Phylogeographic patterns of genetic divergence and speciation in African mole-rats (Family: Bathyergidae). Molecular Ecology 13:613 - 629.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the boostrap. Evolution 39:783 - 791.

Fontanella, F., and M. E. Siddall. 2010. Evaluating hypotheses on the origin and diversification of the ringneck snake *Diadophis punctatus* (Colubridae: Dipsadinae). Zoological Journal of the Linnean Society 158**:**629 - 640.

Fontanillas, P., E. Petit and N. Perrin. 2004. Estimating sex-specific dispersal rates with autosomal markers in hierarchically structured populations. Evolution 58:886 - 894.

Forbes, S. H. and J. T. Hogg. 1999. Assessing populatin structure at high levels of differentiation: microsatellite comparisons of bighorn sheep and large carnivores. Animal Conservation 2:223 - 233.

Forys, E. A. and N. D. Moncrief. 1994. Gene flow among island populations of marsh rice rats (*Oryzomys palustris*). Virginia Journal of Science 45:3 - 11.

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

Friesen, V. L., T. M. Burg and K. D. McCoy. 2007. Mechanisms of population differentiation in seabirds. Molecular Ecology 16:1765 - 1785.

Fulbright, T. E., D. D. Diamond, J. Rappole and J. Norwine. 1990. The Coastal Sand Plain of Southern Texas. Rangelands 12:337 - 340.

Futuyma, D. J. 1998. Evolutionary Biology, Third ed. Sinauer Associates, Inc., Sunderland, MA.

Gaggiotti, O. E., O. Lange, K. Rassmann and C. Gliddons. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Molecular Ecology 8:1513 - 1520.

Gaines, M. S., J. E. Diffendorfer, R. H. Tamarin and T. S. Whittam. 1997. The effects of habitat fragmentation on the genetic structure of small mammal populations. Journal of Heredity 88:294 - 304.

Galbreath, K. E., D. J. Hafner and K. R. Zamudio. 2009. When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American Pika, *Ochotona princeps*). Evolution 63:2848 - 2863.

Galbreath, K. E., D. J. Hafner, K. R. Zamudio and K. Agnew. 2010. Isolation and introgression in the Intermountain West: contrasting gene genealogies revel the complex biogeographic history of the American pika (*Ochotona princeps*). Journal of Biogeography 37:344 - 362.

Gannon, W. L., R. S. Sikes and the Animal Care and Use Committee of the American Society of Mammalogists. 2007. Guidelines of the American Sociey of Mammalogists for the use of wild mammals in research. Journal of Mammalogy 88:809 - 823.

Gardner, L. R. and D. E. Porter. 2001. Stratigraphy and geologic history of a southeastern salt marsh basin, North Inlet, South Carolina, USA. Wetlands Ecology and Management 9:371 - 385.

Garner, A., J. L. Rachlow and J. F. Hicks. 2005. Patterns of genetic diversity and its loss in mammalian populations. Conservation Biology 19:1215 - 1221.

Gates, D. M. 1993. Climate change and its biological consequences. Sinaeuer Associates, Sunderland, MA.

Gauffre, B., E. Petit, S. Brodier, V. Bretagnolle and J. F. Cosson. 2009. Sex-biased dispersal patterns depend on the spatial scale in a social rodent. Proceedings of the Royal Society of London B 276:3487 - 3494.

Geise, L., M. F. Smith and J. L. Patton. 2001. Diversification in the genus *Akodon* (Rodentia: Sigmodontinae) in southeastern South America: mitochondrial DNA sequence analysis. Journal of Mammalogy 82:92 - 101.

Gill, E. B., A. M. Mostrom, and A. L. Mack. 1993. Speciation in North American chickadees. I. Patterns of mtDNA genetic divergence. Evolution 47:195 - 212.

Goldman, E. A. 1918. The rice rats of North America. North American Fauna 43:1 - 100.

Goodyear, N. C. 1987. Distribution and habitat of the silver rice rat, *Oryzomys argentatus*. Journal of Mammalogy 68:692 - 695.

Goodyear, N. C. 1991. Taxonomic status of the silver rice rat, *Oryzomys argentatus*. Journal of Mammalogy 72:723 - 730.

Goodyear, N. C. 1992. Spatial overlap and dietary selection of native rice rats and exotic black rats. Journal of Mammalogy 73:186 - 200.

Google. 2007. GOOGLE EARTH. Ver. 4.1. http://earth.google.com/

Goudet, J. 2001. FSTAT: A program to estimate and test gene diversities and fixation indices. Ver. 2.9.3.

Graham, S. P., D. A. Steen, K. T. Nelson, A. M. Durso and J. C. Maerz. 2010. An overlooked hotspot? Rapid biodiversity assessment reveals a region of exceptional herptofaunal richness in the southeastern United States. Southeastern Naturalist 9:19 - 34.

Greenberg, B. D., J. E. Newbold and A. Sugino. 1983. Intraspecific nucleotide-sequence variablity surrounding the origin of replication in human mitochondrial-DNA. Gene 21:33 - 49.

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

Grill, A., G. Amori, G. Aloise, I. Lisi, et al. 2009. Molecular phylogeography of European *Sciurus vulgaris*: refuge within refugia? Molecular Ecology 18:2687 - 2699.

Hall, E. R. 1960. *Oryzomys couesi* only subspecifically different from the marsh rice rat, *Oryzomys palustris*. Southwestern Naturalist 5:171 - 173.

Hall, E. R. 1981. The mammals of North America. John Wiley & Sons, Inc., New York.

Hall, E. R. and K. R. Kelson. 1959. The mammals of North America. Ronald Press, New York.

Hamilton, W. J. 1946. Habits of the swamp rice rat *Oryzomys palustris*. American Midland Naturalist 36:730 - 736.

Hamilton, W. J. 1955. Two new rice rats (Genus *Oryzomys*) from Florida. Proceedings of the Biological Society Washington 68:83 - 86.

Handley, L. J. L. and N. Perrin. 2007. Advances in our understanding of mammalian sex-biased dispersal. Molecular Ecology 16:1559 - 1578.

Hanson, J. D. and R. D. Bradley. 2008. Molecular diversity within *Melanomys caliginosus* (Rodentia: Oryzomyini): Evidence for multiple species. Occasional Papers, Museum of Texas Tech University 275:1 - 11.

Hanson, J. D., J. L. Indorf, V. J. Swier and R. D. Bradley. 2010. Molecular Divergence within the *Oryzomys palustris* complex: evidence for multiple species. Journal of Mammalogy 91:336 - 347.

Hansson, B. 2010. The use (or misuse) of microsatellite allelic distances in the context of inbreeding and conservation genetics. Molecular Ecology 19:1082 - 1090.

Haq, B. U., W. A. Berggren and J. A. V. Couvering. 1977. Corrected age of the Pliocence/Pleistocene boundary. Nature 269:483 - 488.

Hare, M. P. 2001. Prospects for nuclear gene phylogeography. Trends in Ecology and Evolution 16**:**700 - 706.

Harlan, R. 1837. Description of a new species of quadruped from the order Rodentia inhabiting the United States; by R. Harlan, M. D. *Mus palustris*. American Journal of Science 31:385 - 386.

Harpending, H. 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology 66:591 - 600.

Hartl, D. L. and A. G. Clark. 1997. Principles of Population Genetics. Sinauer Associates, Inc, Sunderland, MA.

Hayes, J. P. and R. G. Harrison. 1992. Variation in mitochondrial DNA and the biogeographic history of woodrats (*Neotoma*) of the eastern United States. Systematic Biology 41:331 - 344.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature 405:907 - 913.

Hewitt, G. M. 2001. Speciation, hybrid zones, and phylogeography - or seeing genes in space and time. Molecular Ecology 10:537 - 549.

Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham, J. B. Johnson, L. Rissler, P. F. Victoriano and A. D. Yoder. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution 54:291 - 301.

Hickerson, M. J. and C. P. Meyer. 2008. Testing comparative phylogeographic models of marine vicariance and dispersal using a hierarchical Bayesian approach. BMC Evolutionary Biology 8:322 - 340.

Hoffman, E. A. and M. S. Blouin. 2004. Evolutionary history of the northern leopard frog: reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA. Evolution 58:145 - 159.

Holsinger, K. E. and B. S. Weir. 2009. Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. Nature Reviews Genetics 10:639 - 650.

Hubisz, M. J., D. Falush, M. Stephens and J. K. Pritchard. 2009. Inferring real population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322 - 1332.

Humphrey, S. R. and H. W. Setzer. 1989. Geographic variation and tazonomic revision of rice rats (*Oryzomys palustris* and *Oryzomys argentatus*) of the United States. Journal of Mammalogy 70:557 - 570.

Hundertmark, K. J., R. T. Bowyer, G. F. Shields and C. C. Schwartz. 2003. Mitochondrial phylogeography of moose (*Alces alces*) in North America. Journal of Mammalogy 84:718 - 728.

Hundertmark, K. J., G. F. Shields, I. G. Udina, T. Bowyer, et al. 2002. Mitochondrial phylogeography of moose (*Alces alces*): Late Pleistocene divergence and population expansion. Molecular Phylogenetics and Evolution 22:375 - 387.

Hürner, H., B. Krystufek, M. Sarà, A. Ribas, et al. 2010. Mitochondrial phylogeography of the edible dormouse (*Glis glis*) in the western Palearctic region. Journal of Mammalogy 91:233 - 242.

Irwin, D. E. 2002. Phylogeographic breaks without geographic barriers to gene flow. Evolution 56:2383 - 2394.

Jackson, N. D. and C. C. Austin. 2009. The combined effects of rivers and refugia generate extreme cryptic fragmentation within the common ground skink (*Scincella lateralis*). Evolution 64:409 - 428.

Kholodova, M. V. 2009. Comparative phylogeography: molecular methods, ecological interpretation. Molecular Biology and Evolution 43:847 - 854.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111 - 120.

Kokko, H. and A. Lopez-Sepulcre. 2006. From individual dispersal to species ranges: perspectives for a changing world. Science 313:789 - 791.

Kruchek, B. L. 2004. Use of tidal marsh and upland habitats by the marsh rice rat (*Oryzomys palustris*). Journal of Mammalogy 85:569 - 575.

Krystufek, B., J. Bryja and E. V. Buzan. 2009a. Mitochondrial phylogeography of the Europen ground squirrel, *Spermophilus citellus*, yields evidence on refugia for steppic taxa in the southern Balkans. Heredity 103:129 - 135.

Krystufek, B., E. V. Buzan, V. Vohralik, R. Zareie and B. Ozkan. 2009b. Mitochondrial cytochrome b sequence yields new insight into the speciation of social voles. Biological Journal of the Linnean Society 98:121 - 128.

Lawson, R. 1987. Molecular studies of thamnophiine snakes. I. The phylogeny of the genus *Nerodia*. Journal of Herpetology 21:140 - 157.

Leigh, D. S. 2008. Late Quaternary climates and river channels of the Atlantic Coastal Plain, Southeastern USA. Geomorphology 101:90 - 108.

Lemmon, E. M., A. R. Lemmon and D. C. Cannatella. 2007. Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudacris*). Evolution 61:2086 - 2103.

Lenormand, T. 2002. Gene flow and the limits to natural selection. Trends in Ecology and Evolution 17:183 - 189.

Li, W. H., M. Gouy, P. M. Sharp, C. O'Huigin and Y. W. Yang. 1990. Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. Proceedings of the National Academy of Sciences of the United States of America 87:6703 - 6707.

Librado, P. and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lidicker, W. Z. 1962. The nature of subspecific boundaries in a desert rodent and its implications for subspecies taxonomy. Systematic Zoology 11:160 - 171.

Liu, F. R., P. E. Moler and M. M. Miyamoto. 2006. Phylogeography of the salamander genus *Pseudobranchus* in the southeastern United States. Molecular Phylogenetics and Evolution 39:149 - 159.

Lowery, G. H. 1974. The mammals of Louisiana and its adjacent waters. Louisiana State University Press, Baton Rouge.

Loxterman, J. L., N. D. Moncrief, R. D. Dueser, C. R. Carlson and J. F. Pagels. 1998. Dispersal abilities and genetic population structure of insular and mainland *Oryzomys palustris* and *Peromyscus leucopus*. Journal of Mammalogy 79:66 – 77.

Luikart, G. and F. W. Allendorf. 1996. Mitochondrial-DNA variation and genetic-population structure in Rocky Mountain bighorn sheep (*Ovis canadensis canadensis*). Journal of Mammalogy 77:109 - 123.

Mace, G. M. and A. Purvis. 2008. Evolutionary biology and practical conservation. Molecular Ecology 17:9 - 19.

Manel, S., M. K. Schwartz, G. Luikart and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution 18: 189 - 197.

Martin, R. A., P. Pelaez-Campomanes, J. G. Honey, D. L. Fox, et al. 2008. Rodent community change at the Pliocene-Pleistocene transition in southwestern Kansas and identification of the *Microtis* immigration event on the Central Great Plains. Palaeogeography, Palaeoclimatology, Palaeoecology 267:196 - 207.

Mayr, E. 1942. Systematics and the Origin of Species from the Viewpoint of a Zoologist. Columbia University Press, New York.

Mayr, E. 1963. Animal Species and Evolution. Harvard University Press, Cambridge, MA.

Mayr, E. 1982. Processes of speciation in animals. *In* C. Barigozzi (ed) Mechanisms of Speciation. Allen R. Liss, New York, pp. 1 - 19.

McGregor, R. L., D. J. Bender and L. Fahrig. 2008. Do small mammals avoid roads because of the traffic? Journal of Applied Ecology 45:117 - 123.

Mendez-Harclerode, F. M., J. D. Hanson, C. F. Fulhorst, M. L. Milazzo, et al. 2005. Genetic diversity within the southern plains woodrat (*Neotoma micropus*) in southern Texas. Journal of Mammalogy 86:180 - 190.

Merriam, C. H. 1901. Synopsis of the rice rats (genus *Oryzomys*) of the United States and Mexico. Proceedings of the Washington Academy of Sciences 3:273 - 295.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718 - 726.

Moran, M. D. 2003. Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos 100:403 - 405.

Morgan, G. S. 1993. Quaternary land vertebrates of Jamaica. *In* R. M. Wright and E. Robinson (eds), Biostratigrapy of Jamaica. Geological Society of America Memoir 182:417 - 442. Boulder, CO: Geological Society of America.

Morgan, G. S. and S. D. Emslie. 2010. Tropical and western influences in vertebrate faunas from the Pliocene and Pleistocene of Florida. Quaternary International 217:143 - 158.

Morin, P. A., J. J. Moore, R. Chakraborty, L. Jin, et al. 1994. Kin selection, social structure, gene flow, and the evolution of Chimpanzees. Science 265:1193 - 1201.

Moritz, C. 1994. Defining "Evolutionary Significant Units" for conservation. Trends in Ecology and Evolution 9:373 - 375.

Moritz, C. 2002. Strategies to protect biological diversity and evolutionary processes that sustain it. Systematic Biology 51:238 - 254.

Moritz, C. and D. P. Faith. 1998. Comparative phylogeography and the identificantion of genetically divergent areas for conservation. Molecular Ecology 7:419 - 429.

Mucci, N., J. Arrendal, H. Ansorge, M. Bailey, et al. 2010. Genetic diversity and landscape genetic structure of otter (*Lutra lutra*) populations in Europe. Conservation Genetics 11:583 - 599.

Musser, G. G. and M. D. Carleton. 1993. Family Muridae. *In* D. E. Wilson and D. M. Reeder (eds), Mammal Species of the World: A Taxonomic and Geographic Reference. Smithsonian Institution Press, Washington, D.C., pp. 501 - 755.

.

Musser, G. G. and M. D. Carleton. 2005. Family Cricetidae. *In* D. E. Wilson and D. M. Reeder (eds), Mammal Species of the World: A Taxonomic and Geographic Reference. Johns Hopkins University Press, Baltimore Maryland, pp. 894 - 1522.

Myers, P., B. Lundrigan and P. K. Tucker. 1995. Molecular phylogenetics of Oryzomine rodents: the genus *Oligoryzomys*. Molecular Phylogenetics and Evolution 4:372 - 382.

Neaves, L. E., K. R. Zenger, R. I. T. Prince, M. D. B. Eldridges and D. W. Cooper. 2009. Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, *Macropus fuliginosus*. Molecular Ecology18:3363 - 3378.

Nichols, R. 2001. Gene trees and species trees are not the same. Trends in Ecology and Evolution 16:358 - 364.

Nordt, L. C., T. W. Boutton, C. T. Hallmark and M. R. Waters. 1994. Late Quaternary vegetation and climate changes in Central Texas based on the isotopic composition of organic carbon. Quaternary Research 41:109 - 120.

Nowak, R. M. 1999. Walker's Mammals of the World. John Hopkins Univeristy Press.

Nyakaana, S., P. Arctander and H. R. Siegismund. 2002. Population structure of the African savannah elephant inferred from mitochondrial control region sequences and nuclear microsatellie loci. Heredity 89:90 - 98.

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

Odum, E. P. 2002. The southeastern region: a biodiveristy haven for naturalists and ecologists. Southeastern Naturalist 1:1 - 12.

Osentoski, M. F. and T. Lamb. 1995. Intraspecific phylogeography of the gopher tortoise, *Gopherus polyphemus*: RFLP analysis of amplified mtDNA segments. Molecular Ecology 4:709 - 718.

Paradiso, J. L. 1960. Size variation in the rice rat. Journal of Mammalogy 41:516 - 517.

Pauly, G. B., O. Piskurek, and H. B. Shaffer. 2007. Phylogeographic concordance in the southeastern United States: the flatwoods salamander, *Ambystoma cingulatum*, as a test case. Molecular Ecology 16:415 - 429.

Peakall, R., M. Ruibal and D. B. Lindenmayer. 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. Evolution 57:1182 - 1195.

Peppers, L. L. and R. D. Bradley. 2000. Cryptic species in *Sigmodon hispidus*: evidence from DNA sequences. Journal of Mammalogy 81:332 - 343.

Phillips, C. A. 1994. Geographic distribution of mitochondrial DNA variants and the historical biogeography of the spotted salamander, *Ambystoma maculatum*. Evolution 48:597 - 607.

Phillips, C. D., C. A. Henard, and R.S. Pfau. 2007. Amplified fragment length polymorphism and mitochondrial DNA analyses reveal patterns of divergence and hybridization in the hispid cotton rat (*Sigmodon hispidus*). Journal of Mammalogy 88:351 - 359.

Piry, S., G. Luikart and J. M. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90:502 - 503.

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

Pritchard, J. K. and W. Wen. 2003. Documentation for STRUCTURE Software: Version 2. Available at: http://pritch.bsd.uchicago.edu.

Rambaut, A. 2006. FIGTREE: Tree Figure Drawing Tool. Ver. 1.3.1. Institute of Evolutionary Biology, University of Edinburgh.

Rambaut, A. and A. J. Drummond. 2007. TRACER v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Ramey, R. R. 1995. Mitochondrial DNA variation, population structure, and evolution of mountain sheep in the south-western United States and Mexico. Molecular Ecology 4:429 - 439.

Ray, C. E. 1967. Pleistocene mammals of Ladds, Barrow County, Georgia. Bulletin of the Georgia Academy of Science 25:120 - 150.

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

Richards, R. L. 1979. Rice rat (*Oryzomys palustris*) remains from southern Indiana caves. Proceedings of the Indiana Academy of Science 89:425 - 431.

Richardson, C. J. 2009. The Everglades: North America's subtropical wetland. Wetlands Ecology and Management. Published online, DOI 10.1007/s11273-009-9156-4.

Riddle, B. R. 1996. The molecular phylogeographic bridge between deep and shallow history in continental biotas. Trends in Ecology and Evolution 11:207 - 211.

Riddle, B. R., M. N. Dawson, E. A. Hadly, D. J. Hafner, et al. 2008. The role of molecular genetics in sculpting the future of integrative biogeography. Progress in Physical Geography 32:173 - 202.

Riddle, B. R., D. J. Hafner and L. F. Alexander. 2000. Phylogeography and systematics of the *Peromyscus eremicus* species group and the historical biogeography of the North American warm regional deserts. Molecular Phylogenetics and Evolution 17:145 - 160.

Riddle, B. R., R. L. Honeycutt and P. L. Lee. 1993. Mitochondrial-DNA phylogeography in northern grasshopper mice (*Onychomys leucogaster*) - the influence of Quaternary climatic oscillations on population dispersion and divergence. Molecular Ecology 2:183 - 193.

Rogers, A. R. 1995. Genetic evidence for a Pleistocene population explosion. Evolution 49:608 - 615.

Rogers, A. R. and H. Harpending. 1992. Populations growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution 9:552 - 569.

Rohlf, F. J. 1973. ALGORITHM 76. Hierarchical clustering using the minimum spanning tree. Computer Journal 6:93 - 95.

Ronquist, F. and J. P. Huelsenbeck. 2003. MRBAYES 3: Bayesian pylogenetic inference under mixed models. Bioinformatics 19:1572 - 1574

Rowe, K. C., E. J. Heske and K. N. Paige. 2006. Comparative phylogeography of eastern chipmunks and white-footed mice in relation to the individualistic nature of species. Molecular Ecology 15:4003 - 4020.

Rundle, H. D. and P. Nosil. 2005. Ecological speciation. Ecology Letters 8:336 - 352.

Ryder, O. A. 1986. Speices conservation and systematics: the dilemma of subspecies. Trends in Ecology and Evolution 1:9 - 10.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, et al. 1988. Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. Science 239:487 - 491.

Saldanha, A. J. 2004. JAVA TREEVIEW - extensible visualization of microarray data. Bioinformatics 20:3246 - 3248.

Santucci, F., B. C. Emeron and G. M. Hewitt. 1998. Mitochondrial DNA phylogeography of European hedgehogs. Molecular Ecology 7:1163 - 1172.

Sbisà, E., F. Tanzariello, A. Reyes, G. Pesole and C. Saccone. 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205:125 - 140.

Schimdt, C. A. and M. D. Engstrom. 1994. Genic variation and systematics of rice rats (*Oryzomys palustris* species group) in southern Texas and northeastern Tamaulipas, Mexico. Journal of Mammalogy 75:914 - 928.

Schipper, J., J. S. Chanson, F. Chiozza, N. A. Cox, et al. 2008. The status of the world's land and marine mammals: diversity, threat, and knowledge. Science 322:225 - 230.

Schneider, C. J., M. Cunningham and C. Moritz. 1998. Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. Molecular Ecology 7:487 - 498.

Schneider, N., L. Chikhi, M. Currat and U. Radespiel. 2010. Signals of recent spatial expansions in the grey mouse lemur (Microcebus murinus). BMC Evolutionary Biology 10:105 - 122.

Scott, J. M., T. Loveland, K. Gergely, J. Strittholt and N. Staus. 2004. National Wildlife Refuge System: Ecological context and integrity. Natural Resources Journal 44:1041 - 1066.

Selkoe, K. A. and R. J. Toonen. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters 9:615 - 629.

Shipp-Pennock, M. A., W. D. Webster and D. W. Freshwater. 2005. Systematics of the white-footed mouse (*Peromyscus leucopus*) in the mid-Atlantic region. Journal of Mammalogy 86:803 - 813.

Simmons, M. P. and H. Ochoterena. 2000. Gaps as characters in sequenced-based phylogenetic analyses. Systematic Biology 49:369 - 381.

Simonsen, B., H. Siegismund and P. Arctander. 1998. Population structure of African buffalo inferred from mtDNA sequences and microsatellite loci: high variation but low differentiation. Molecular Ecology 7:225 - 237.

Sites, J. W. and J. C. Marshall. 2003. Delimiting species: a Renaissance issue in systematic biology. Trends in Ecology and Evolution 18:462 - 469.

Slatkin, M. 1991. Inbreeding coefficients and coalescence times. Genetical Research 58:167 - 175.

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

Slatkin, M. and R. R. Hudson. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555 - 562.

Smith, A. T. and M. L. Weston. 1990. *Ochotona princeps*. Mammalian Species 352:1 - 8.

Smith, M. F. and J. L. Patton. 1993. The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe. Biological Journal of the Linnean Society 50:149 - 177.

Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology 15:4261 - 4293.

Spitzer, N. C. 1983. Aspects of the biology of the silver rice rat, *Oryzomys argentatus*. M. S. thesis, University of Rhode Island, Kingston, 100 pp.

Spitzer, N. C. and J. J. D. Lazell. 1978. A new rice rat (Genus *Oryzomys*) from Florida's Lower Keys. Journal of Mammalogy 59:787 - 792.

Stamatakis, A. 2006. RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of taxa and Mixed Models. Bioinformatics 22:2688 - 2690.

Stamatakis, A., P. Hoover and J. Rougemont. *In preparation*. A rapid bootstrap algorithm for the RAxML Web-Servers.

Steele, M. A. 1998. *Tamiasciurus hudsonicus*. Mammalian Species 586:1 - 9.

Stewart, J. R., A. M. Lister, I. Barnes and L. Dalén. 2010. Refugia revisited: individualistic responses of species in space and time. Proceedings of the Royal Society of London B 277:661 - 671.

Swenson, N. G. and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. The American Naturalist 166:581 - 591.

Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Ver. 4.0b10. Sinauer Associates.

Taberlet, P., L. Fumagalli, A-G. Wust-Saucy and J-F. Cosson. 1998. Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology 7:453 - 464.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437 - 460.

Tajima, F. 1989a. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585 - 596.

Tajima, F. 1989b. The effect of change in population size on DNA polymorphism. Genetics 123:597 - 601.

Tajima, F. 1993. Measurement of DNA polymorphism. *In* N. Takahata and A. G. Clark (eds), Mechanisms of Molecular Evolution. Japan Scientific Societies Press, Sinauer Associates, Inc., Tokyo, Sunderland, MA, pp. 37 - 59.

Tamura, K., J. Dudley, M. Nei and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetic Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596 - 1599.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. American Mathematical Society: Lectures on Mathematics in the Life Sciences 17:57 - 86.

Thomas, O. 1893. Notes on some Mexican *Oryzomys*. Annals and Magazine of Natural History Series 6:402 - 405.

Thomas, O. 1898. On indigenous Muridae in the West Indies; with the description of a new Mexican *Oryzomys*. Annals and Magazine of Natural History Series 7:176 - 180.

Toonen, R. and S. Hughes. 2001. Increased throughput for fragment analysis on ABI Prism 377 automated sequencer using a membrane comb and STRand software. Biotechniques 31:1320 - 1324.

U.S. Fish and Wildlife Service. 1999. Rice rat *Oryzomys palustris natator*. In: Multi-Species Recovery Plan for South Florida. Pp. 173 - 186.

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

Van Oosterhout, C., D. Weetman and W. F. Hutchinson. 2006. Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. Molecular Ecology Notes 6:255 - 256.

Van Vuuren, B. J. and T. J. Robinson. 1997. Genetic population structure in the yellow mongoose, *Cynictis penicillata*. Molecular Ecology 6:1147 - 1153.

Vega, R., E. Vazquez-Dominguez, A. Mejia-Puente and A. D. Cuaron. 2007. Unexpected high levels of genetic variability and the population structure of an island endemic rodent (*Oryzomys couesi cozumelae*). Biological Conservation 137:210 - 222.

Vilà, C., I. R. Amorim, J. A. Leonard, D. Posada, et al. 1999. Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. Molecular Ecology 8:2089 - 2103.

Vogler, A. P. and R. DeSalle. 1993. Pylogeographic patterns in coastal North American tiger beetles (*Cicindela dorsalis* Say) inferred from mitochondrial DNA sequences. Evolution 47:1192 - 1202.

Walker, D., V. J. Burke, I. Barak and J. C. Avise. 1995. A comparison of mtDNA restriction sites vs. control region sequences in phylogeographic assessment of the musk turtle (*Sternotherus minor*). Molecular Ecology 4:365 - 373.

Wang, Y. Q., C. R. Hughes, E. A. Gines-Candelaria and M. S. Gaines. 2000. Polymorphic microsatellite loci of *Oryzomys palustris*, the marsh rice rat, in South Florida detected by silver staining. Molecular Ecology 9:1931 - 1932.

Wang, Y., D. A. Williams and M. S. Gaines. 2005. Evidence for a recent genetic bottleneck in the endangered Florida Keys silver rice rat (*Oryzomys argentatus*) revealed by microsatellite DNA analyses. Conservation Genetics 6:575 - 585.

Webb, S. D. 1974. Pleistocene Mammals of Florida. University Press of Florida, Gainesville.

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

Weiss, S. and N. Ferrand (eds). 2007. Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of Europen Biodiversity. Springer, Dordecht.

Weksler, M. 2003. Phylogeny of Neotropical Oryzomyine rodents (Muridae: Sigmodontinae) based on the nuclear IRBP exon. Molecular Phylogenetics and Evolution 29:331 - 349.

Weksler, M. 2006. Phylogenetic relationships of Oryzomine rodents (Muroidea: Sigmodontinae): separate and combined analyses of morphological and molecular data. Bulletin of the American Museum of Natural History 296:1 - 149.

Weksler, M., A. R. Percequillo and R. S. Voss. 2006. Ten new genera of Oryzomyine rodents (Cricetidae: Sigmodontinae). American Museum Novitates 3537.

Wertheim, J. O., M. J. Sanderson, M. Worobet and A. Jjork. 2010. Relaxed molecular clocks, the bias-variance trade-off, and the quality of phylogenetic inference. Systematic Biology 59:1 - 8.

Wilson, G. M., R. A. DenBussche, K. McBee, L. A. Johnson and C. A. Jones. 2005. Intraspecific phylogeography of red squirrels (*Tamiasciurus hudsonicus*) in the Central Rocky Mountain region of North America. Genetica 125:141 - 154.

Wolfe, J. L. 1982. *Oryzomys palustris*. Mammalian Species 176:1 - 5.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics 15:323 - 354.

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

Wright, S. 1978. Evolution and the Genetics of Population, Variability Within and Among Natural Populations. The University of Chicago Press, Chicago.

Yang, D-S. and G. J. Kenagy. 2009. Nuclear and mitochondrial DNA reveal contrasting evolutionary processes in populations of deer mice (*Peromyscus maniculatus*). Molecular Ecology 18:5115 - 5125.

Zamudio, K. R. and W. K. Savage. 2003. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). Evolution 57:1631 - 1652.

Zenger, K. R., M. D. B. Eldridge and D. W. Cooper. 2003. Intraspecific variation, sex-biased dispersal and phylogeography of the eastern grey kangaroo (*Macropus giganteus*). Heredity 91:153 - 162.

Appendix A: Specimens Examined

All specimens included in this study are listed below by species, subspecies and geographic locality. Identification numbers for samples and Genbank accession numbers for Cytb only (if available) are given in parentheses. Museum and collection abbreviations are: Angelo State Natural History Collections (ASNHC, ASK), Richard Stevens – Louisiana State University, Natural History Museum (LSUM), Museum of Southwestern Biology (MSB, NK), Museum of Texas Tech University (TTU, TK), Robert K. Rose – Old Dominion University (SCVA, VNWR), Kent Edmonds – Indiana University Southeast (DEL), David Webster – University of North Carolina Wilmington (UNCW), Magaly Massennet and Emily Woods – J. N. "Ding" Darling National Wildlife Refuge, Florida (SCVA), Amanda Crouse (Keys), Phillip Frank (SK), Jane Indorf – University of Miami, Florida (EVGL, KPSP, LPI, SJBP, ANERR).

Oryzomys couesi – Mexico: Veracruz; Estacion Biologia Morr. (MSB75575/NK27005); Rancho el Quetzal (MSB75570/NK27111); Texas: Cameron County; Los Palomas Wildlife Management Area, Resaca de la Palma Unit (TK72660/TTU77220, DQ370034; TTU77221/TK72661, EU074662); Port Isabel (ASNHC3014/ASK2255, identified as *O. palustris*).

Oryzomys couesi couesi – Honduras: Atlantida; Jardin Botanico Lancetilla (TTU103830/TK136208, EU074666); Olancho; 4 km E Catacamas, Escuela de Sembrador (TTU84697/TK102040, DQ185383)

Oryzomys mexicanus mexicanus – Mexico: Oaxaca; Las Minas (TK93218, DQ185385; TTU82862/TK93244, DQ185386)

Oryzomys palustris argentatus (formerly assigned to *O. p. natator* or *O. argentatus*) – Florida; Monroe County; National Key Deer Refuge, Lower Florida Keys (Keys431, Keys10404, Keys10412, Keys10414, Keys10415, Keys10417, Keys10418, Keys10419, Keys10429, Keys10431), Great White Heron National Wildlife Refuge (SK4, SK5, SK15).

Oryzomys palustris palustris (formerly assigned to *O. p. coloratus*) – Florida; Miami-Dade County, Everglades National Park, near Homestead, Rock Reef Pass (EVGL01, FJ974114; EVGL02, FJ974115; EVGL03; EVGL04; EVGL05, GQ148811; EVGL06, EU074639; EVGL07; EVGL08; EVGL09; EVGL10; EVGL11; EVGL12; EVGL13; EVGL14; EVGL15; EVGL16; EVGL17; EVGL18; EVGL19; EVGL20).

Oryzomys palustris palustris (formerly assigned to *O. p. natator*) – Florida; Okeechobee County, Kissimmee Prairie Preserve State Park (KPSP01, FJ974108; KPSP02, FJ974109); St Johns County; Rattlesnake Island, Fort Matanzas (UNCW12881; UNCW12870).

Oryzomys palustris palustris – Alabama; Colbert County, Cherokee, Natchez Trace Parkway Mile Marker 312.4 (MSB81541/NK52123; MSB81542/NK52117; MSB81543/NK52114, EU74636); Tallapoosa County, Horseshoe Bend National Military

Park (MSB81643, FJ974112; MSB81644, FJ74113); Delaware; lab colony descended from individuals from Sussex County, Lewes, (DEL01, DEL02); Florida; Franklin County; Apalachicola Bay, Apalachicola National Estuarine Research Reserve (ANERR01, FJ974106; ANERR02, FJ974107; ANERR03); Gulf County, Port St. Joe, St. Joseph Bay State Buffer Preserve (SJBP01, FJ974110; SJBP02, FJ974111); Georgia; Glynn County, Fort Frederica (UNCW11027); Gwinnett County, Suwanee (UNCW10075, UNCW10076); North Carolina; Dare County, Pea Island (UNCW2569); New Hanover County (UNCW2555, UNCW2557, UNCW2215, UNCW19444, UNCW19440, UNCW19441, UNCW19449, UNCW19443, UNCW19438, UNCW19442, UNCW19446, UNCW19445, UNCW19448, UNCW19437, UNCW19439, UNCW19447); South Carolina; Richland County, Congaree Swamp National Monument (MSB74956, EU074637); Virginia; Norfolk County, 0.8 km E US Hwy 17 or 9.7 km from Virginia/North Carolina border (SCVA01; SCVA02; SCVA03; SCVA04; SCVA05; SCVA06; SCVA07; SCVA08; SCVA09; SCVA10; SCVA11; SCVA12; SCVA13; SCVA14; SCVA15, EU074640; SCVA16; SCVA17; SCVA18); Northampton County, Virginia National Wildlife Refuge (VNWR01, VNWR02, VNWR03, VNWR04, VNWR05)

Oryzomys palustris palustris (formerly assigned to *O. p. planirostris*) – Florida; Lee County, Little Pine Island (LPI01, FJ974116; LPI02, FJ974117; LPI03, LPI04, LPI05, LPI06, LPI07, LPI08)

Oryzomys palustris sanibeli – Florida; Lee County; Sanibel Island, J. N. "Ding" Darling National Wildlife Refuge (SBI01, FJ974118; SBI02, EU074638; SBI03, FJ974119; SBI04, SBI05, SBI06, SBI07, SBI08, SBI09, SBI10, SBI11, SBI12).

Oryzomys texensis (formerly assigned to *O. p. texensis*) – Arkansas; Crittenden County, West Memphis (TTU82963, FJ74129); Louisana; Cameron Parish, Rockefeller Refuge (LSUM8428, FJ974123; LSUM8433, FJ974124; LSUM8436, FJ974125; LSUM8432, LSUM8438, LSUM8439, LSUM8441, LSUM8446, LSUM8447, LSUM8448, LSUM8449, LSUM8450, LSUM8451; LSUM8452, LSUM8453, LSUM8454, LSUM8455, LSUM8460, LSUM8437, LSUM8443); Mississippi; Lee County, Tupelo, Natchez Trace Parkway Mile Marker 261.8 (MSB81544, EU074643); Oklahoma; Okmulgee County; 4.8 km E Dewar, Eufaula Wildlife Management Area (TTU62980/TK27995, DQ37032; TTU62979/TK27994, TTU62978/TK28376, TTU62981/TK28375); Tennessee; Shelby County; 8.2 km N Memphis on HWY 388 (TTU79152/TK74922, FJ974126; TTU79153/TK74923, FJ974127); Edward J. Meeman Biological Station (TTU79154/TK83601, FJ974128); Texas; Anderson County; Gus Engeling Wildlife Management Area (TTU75425/TK52111, TTU97951/TK92582); Brazoria County; Peach Point Wildlife Management Area (TTU71591/TK51538, TTU71592/TK51536, TTU71593/TK51539, TTU71594/TK51540, TTU71595/TK51629, TTU71590/TK51541, TTU71559/TK51567, TTU71560/TK51568, TTU71561/TK51569, TTU71564/TK51570, TTU71558/TK51571, TTU71563/TK51572, TTU71583/TK51602, TTU71582/TK51605, TTU71568/TK51612, TTU71562/TK51613, TTU71567/TK51614, TTU78331/TK53661, TTU78332/TK53670); Calhoun County; Guadalupe Delta Wildlife

Management Area (TTU75163/TK51652, TTU75191/TK51673, TTU75192/TK51676, TTU75193/TK51744; TTU75177/TK51628, DQ370031); Cameron County; Port Isabel (ASNHC2894/ASK775, ASNHC2900/ASK781, ASNHC2934/ASK815, ASNHC2915/ASK796, ASNHC2923/ASK804, ASNHC2935/ASK816, ASNHC2997/ASK2237, ASNHC2983/ASK832, ASNHC2917/ASK798, ASNHC3001/ASK2241, ASNHC2980/ASK829, ASNHC2922/ASK803, ASNHC2886/ASK767, ASNHC2908/ASK789, ASNHC3024/ASK905, ASNHC2977/ASK826, ASNHC2890/ASK771); Freestone County; Richland Creek Wildlife Management Area (TTU75268/TK52124, TTU75307/TK52154, TTU75308/TK52164, TTU75309/TK52163, TTU75310/TK52121; TTU74311/TK52128, EU074642; TTU75312/TK52166); Galveston County; Texas City, Virginia Point (TTU82870/TK90141, TTU82871/TK90142, TTU82872/TK90143, TTU82873/TK90144, TTU82874/TK90145, TTU82875/TK90146, TTU82876/TK90147, TTU82878/TK90149, TTU82879/TK90150, TTU82880/TK90151, TTU82881/TK90152, TTU82882/TK90153, TTU82883/TK90154, TTU82884/TK90155, TTU82885/TK90156, TTU82888/TK90159, TTU82890/TK90161, TTU82891/TK90162, TTU82892/TK90163, TTU82869/TK90140); San Patricio County; Odem (ASNHC3035/ASK2343, ASNHC3037/ASK2345, ASNHC3043/ASK2351; ASNHC3036/ASK2344; ASNHC3040/ASK2348; ASNHC 3034/ASK2342; ASNHC3038/ASK2346; ASNHC3032/ASK2340; ASNHC3044/ASK 2352; ASNHC 3033/ASK 2341; ASNHC3031/ASK 2339; ASNHC 3039/ASK2347; ASNH3042/ASK2350); Willacy County; Port Mansfield (ASNHC3057/ASK982, ASNHC3049/ASK977, ASNHC3054/ASK979, ASNHC3056/ASK981, ASNHC3055/ASK980,

ASNHC3051/ASK994, ASNHC3052/ASK993, ASNHC3048/ASK976, ASNHC 3047/ASK975, ASNHC3050/ASK978, ASNHC3045/ASK970); Mexico; Tamaulipas, Matamoros (ASNHC3432/ASK2206, FJ974120; ASNHC3420/ASK2207, ASNHC3434/ASK2208, ASNHC3437/ASK2211, ASNHC3438/ASK2212; ASNHC3439/ASK2213, FJ974122; ASNHC3440/ASK2214, ASNHC3442/ASK2216, FJ974121/xxxx, ASNHC3420/ASK2218, ASNHC3421/ASK2219, ASNHC3422/ASK2220, ASNHC3423/ASK2221, ASNHC3424/ASK2222, ASNHC3425/ASK2223, ASNHC3428/ASK2226, ASNHC3430/ASK2228, ASNHC3444/ASK2230, ASNHC3446/ASK2232, ASNHC3447/ASK2233, ASNHC3448/ASK2234)

Appendix B: Microsatellite Allele Frequencies at Nine Loci in 12 Populations of the Marsh Rice Rat (*Oryzomys palustris*).

Population abbreviations are EFL (Everglades, Miami-Dade County, Florida), NVA (Southern Chesapeake Bay, Norfolk County, Virginia), SFL (Sanibel Island, Lee County, Florida), CLA (Cameron Parish, Louisiana), BTX (Brazoria County, Texas), KFL (Lower Keys, Monroe County, Florida), TMX (Matamoros, Tamaulipas, Mexico), CMT (Cameron County, Texas), WTX (Willacy County, Texas), STX (San Patricio County, NNC (New Hanover County, North Carolina), Texas), and GTX (Galveston County, Texas).

Allele	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
Locus AAT03												
107	0	0	0	0	0	0	0.025	0	0	0	0.10526	0
110	0	0	0	0	0	0	0.35	0.20588	0.90909	0	0.10526	0
113	0.025	0	0	0	0	0	0	0	0	0	0	0
116	0	0	0	0	0	0.11538	0	0	0	0	0	0.05263
119	0.025	0	0	0	0.10526	0.34615	0	0.02941	0	0	0.05263	0
122	0.025	0	0	0.025	0	0.03846	0	0.17647	0	0	0	0
125	0.175	0.11765	0	0.15	0.13158	0	0	0.17647	0	0	0.13158	0.28947
128	0.175	0.08824	0.29167	0.25	0.28947	0.15385	0.05	0.11765	0.09091	0.10714	0.05263	0.02632
131	0.175	0.38235	0.375	0.15	0.13158	0.03846	0.1	0.05882	0	0.25	0.26316	0.39474
134	0.15	0.29412	0.33333	0.225	0.05263	0.26923	0.025	0	0	0.17857	0.13158	0.13158
137	0.125	0.05882	0	0.125	0.05263	0.03846	0.3	0.02941	0	0.28571	0.13158	0.07895
140	0.1	0.05882	0	0.025	0.10526	0	0.15	0.11765	0	0.10714	0	0

143	0.025	0	0	0	0.07895	0	0	0.08824	0	0.07143	0	0
146	0	0	0	0.05	0.02632	0	0	0	0	0	0.02632	0
149	0	0	0	0	0.02632	0	0	0	0	0	0	0
86	0	0	0	0	0	0	0	0	0	0	0	0.02632

Locus												
AAT10	**EFL**	**NVA**	**SBI**	**CLA**	**BTX**	**KFL**	**TMX**	**CMT**	**WTX**	**STX**	**NNC**	**GTX**
118	0	0	0.45833	0	0	0	0.05263	0	0	0	0	0
121	0.025	0	0.08333	0	0	0	0	0	0	0.03571	0	0
124	0.075	0.08824	0	0	0	0	0.42105	0.08824	0.27273	0	0.03333	0
127	0.275	0	0.125	0.025	0.02632	0	0.07895	0	0	0.21429	0	0
130	0.125	0.11765	0.16667	0.275	0.07895	0.03846	0.02632	0.32353	0.45455	0.28571	0.06667	0.13158
133	0.025	0.5	0.125	0.075	0.05263	0	0.05263	0	0.04545	0.03571	0.13333	0.07895
136	0.15	0.20588	0	0.075	0.15789	0.57692	0	0	0	0	0.26667	0.10526
139	0	0.05882	0	0.2	0.18421	0.19231	0.31579	0.17647	0.18182	0.25	0.16667	0.02632
142	0.15	0.02941	0	0.15	0.18421	0.03846	0.02632	0.02941	0	0.07143	0.06667	0.39474
145	0.075	0	0	0.1	0.07895	0.15385	0	0.02941	0.04545	0	0.13333	0.07895
148	0.075	0	0	0.075	0.02632	0	0	0	0	0.07143	0	0.15789
151	0	0	0	0.025	0	0	0	0	0	0.03571	0.03333	0
154	0.025	0	0	0	0.10526	0	0.02632	0.35294	0	0	0	0
157	0	0	0.04167	0	0.05263	0	0	0	0	0	0.1	0.02632
160	0	0	0	0	0.02632	0	0	0	0	0	0	0
163	0	0	0	0	0.02632	0	0	0	0	0	0	0

Locus AAT16	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
100	0.025	0	0	0	0	0	0.025	0.29412	0	0.42857	0.11765	0.10526
103	0.075	0.02941	0.125	0.05	0.13158	0	0	0	0	0	0.14706	0.23684
106	0.2	0.11765	0.58333	0.125	0.05263	0.73077	0	0	0	0.03571	0.08824	0.10526
109	0.175	0.44118	0.08333	0.15	0.23684	0	0	0.17647	0	0.28571	0.17647	0.10526
112	0.15	0	0.04167	0.125	0.18421	0	0.5	0.35294	0.68182	0.10714	0.05882	0.10526
115	0.1	0.08824	0.125	0.325	0.28947	0.03846	0.35	0.17647	0.31818	0.07143	0.08824	0.13158
118	0.075	0.02941	0.04167	0.1	0.07895	0	0.1	0	0	0.07143	0.08824	0.07895
121	0	0.08824	0	0.05	0	0	0.025	0	0	0	0	0
124	0	0	0	0	0	0	0	0	0	0	0	0.02632
94	0.1	0.20588	0	0.075	0	0.23077	0	0	0	0	0.23529	0
97	0.1	0	0	0	0.02632	0	0	0	0	0	0	0.10526

Locus AAT21	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
148	0	0.02941	0	0	0	0	0.025	0.02941	0	0	0.02632	0
157	0	0	0	0	0	0	0	0	0	0.03571	0	0
160	0.075	0.20588	0	0.125	0.02632	0	0.275	0.14706	0.13636	0	0.21053	0.21053
163	0.05	0.02941	0	0	0	0	0	0	0	0.07143	0.10526	0
166	0.025	0.08824	0	0.025	0.07895	0	0	0.05882	0.04545	0.07143	0.05263	0
169	0.05	0	0.04167	0.1	0.18421	0.15385	0	0.29412	0	0.21429	0.05263	0
172	0.175	0.05882	0.375	0.2	0.13158	0	0	0	0	0.03571	0.02632	0.07895
175	0.175	0.08824	0	0.1	0.10526	0.15385	0	0.02941	0.13636	0.03571	0.02632	0.13158
178	0.15	0.11765	0.20833	0.15	0.15789	0.19231	0.35	0.17647	0.54545	0.10714	0.21053	0.26316

179	0	0	0	0	0	0	0.025	0	0	0	0	0
181	0.125	0.20588	0.125	0.075	0.15789	0.19231	0.125	0.23529	0.13636	0.28571	0.15789	0.15789
184	0.15	0	0.08333	0.075	0.13158	0.30769	0.125	0.02941	0	0	0.05263	0.10526
187	0.025	0.17647	0.16667	0.075	0.02632	0	0	0	0	0.10714	0.02632	0.05263
190	0	0	0	0.075	0	0	0.025	0	0	0.03571	0	0
193	0	0	0	0	0	0	0.05	0	0	0	0	0
196	0	0	0	0	0	0	0	0	0	0	0.05263	0

Locus AAT26	**EFL**	**NVA**	**SBI**	**CLA**	**BTX**	**KFL**	**TMX**	**CMT**	**WTX**	**STX**	**NNC**	**GTX**
102	0.275	0	0.125	0	0	0	0	0	0	0.17857	0	0
105	0	0	0	0	0.05263	0	0	0	0.05	0.07143	0	0
108	0.05	0	0.29167	0	0	0	0	0	0	0	0	0
111	0.025	0	0	0.025	0	0.11538	0	0	0	0	0	0
114	0.025	0.02941	0	0.025	0.05263	0.73077	0	0	0	0	0.03846	0.05263
117	0.05	0.05882	0	0	0.05263	0.03846	0	0.11765	0	0.28571	0	0.10526
120	0.05	0.23529	0	0.1	0	0.03846	0.05556	0.02941	0.15	0.14286	0.19231	0
123	0.125	0.32353	0	0.175	0.42105	0	0.16667	0.5	0.1	0.25	0.07692	0.36842
126	0.125	0.05882	0.20833	0.325	0.15789	0.07692	0.19444	0.23529	0.05	0	0.19231	0.15789
129	0.075	0.11765	0	0.25	0.15789	0	0.38889	0.05882	0.4	0.03571	0.26923	0.07895
132	0.1	0.05882	0	0.1	0.05263	0	0.19444	0.02941	0.25	0.03571	0.07692	0.10526
135	0.05	0.11765	0.16667	0	0.05263	0	0	0.02941	0	0	0.11538	0.05263
138	0.025	0	0.16667	0	0	0	0	0	0	0	0	0.05263
141	0.025	0	0	0	0	0	0	0	0	0	0	0.02632
144	0	0	0.04167	0	0	0	0	0	0	0	0	0
192	0	0	0	0	0	0	0	0	0	0	0.03846	0

Locus AAT28	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
102	0.275	0	0.125	0	0	0	0	0	0	0.17857	0	0
105	0	0	0	0	0.05263	0	0	0	0.05	0.07143	0	0
108	0.05	0	0.29167	0	0	0	0	0	0	0	0	0
111	0.025	0	0	0.025	0	0.11538	0	0	0	0	0	0
114	0.025	0.02941	0	0.025	0.05263	0.73077	0	0	0	0	0.03846	0.05263
117	0.05	0.05882	0	0	0.05263	0.03846	0	0.11765	0	0.28571	0	0.10526
120	0.05	0.23529	0	0.1	0	0.03846	0.05556	0.02941	0.15	0.14286	0.19231	0
123	0.125	0.32353	0	0.175	0.42105	0	0.16667	0.5	0.1	0.25	0.07692	0.36842
126	0.125	0.05882	0.20833	0.325	0.15789	0.07692	0.19444	0.23529	0.05	0	0.19231	0.15789
129	0.075	0.11765	0	0.25	0.15789	0	0.38889	0.05882	0.4	0.03571	0.26923	0.07895
132	0.1	0.05882	0	0.1	0.05263	0	0.19444	0.02941	0.25	0.03571	0.07692	0.10526
135	0.05	0.11765	0.16667	0	0.05263	0	0	0.02941	0	0	0.11538	0.05263
138	0.025	0	0.16667	0	0	0	0	0	0	0	0	0.05263
141	0.025	0	0	0	0	0	0	0	0	0	0	0.02632
144	0	0	0.04167	0	0	0	0	0	0	0	0	0
192	0	0	0	0	0	0	0	0	0	0	0.03846	0

Locus AAT40	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
124	0	0	0	0	0.02632	0	0	0	0	0.10714	0	0.02632
127	0	0	0	0.025	0	0	0.325	0.29412	0.40909	0.07143	0	0
130	0	0	0	0	0.15789	0	0	0	0	0.03571	0	0
133	0	0.05882	0.125	0.075	0	0	0	0	0	0	0.05263	0.13158
136	0.2	0	0.45833	0.175	0.05263	0	0	0.23529	0.09091	0.03571	0.07895	0.13158
139	0.2	0.17647	0.20833	0.1	0.05263	0	0	0	0	0.03571	0.13158	0.13158
142	0.25	0.29412	0.125	0.175	0.26316	0.03846	0	0	0.04545	0.10714	0.21053	0.21053
145	0.125	0.32353	0	0.1	0.15789	0.38462	0.475	0.29412	0.36364	0.14286	0.21053	0.07895
148	0.05	0.11765	0.04167	0.2	0.13158	0.38462	0.025	0.02941	0.04545	0.21429	0.13158	0.02632
151	0	0	0.04167	0.125	0.07895	0.07692	0.175	0.05882	0	0.17857	0.13158	0.07895
154	0.1	0	0	0	0.02632	0	0	0.08824	0.04545	0.07143	0.05263	0.10526
157	0.075	0	0	0.025	0.05263	0.11538	0	0	0	0	0	0.05263
160	0	0.02941	0	0	0	0	0	0	0	0	0	0.02632

Locus AAT60	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
118	0	0	0	0	0	0	0.025	0.02941	0	0	0	0
124	0	0	0	0	0.02632	0	0.5	0.20588	0.36364	0.03571	0	0
127	0	0	0	0	0	0	0	0	0	0.07143	0	0
130	0.05	0	0	0	0	0	0	0	0	0	0	0.02632
133	0.05	0	0	0.025	0.07895	0	0	0.05882	0.13636	0.07143	0.05882	0
136	0	0.02941	0	0.075	0	0	0	0	0	0.10714	0.02941	0.02632
139	0.05	0.29412	0	0.05	0.10526	0	0.4	0.32353	0.40909	0	0.08824	0.02632
142	0.1	0.20588	0.625	0.05	0	0	0	0.11765	0	0.10714	0.14706	0.02632
145	0	0	0	0.025	0.07895	0	0.025	0.08824	0.04545	0.03571	0.11765	0.02632
148	0.075	0.05882	0.29167	0.125	0.13158	0	0	0	0	0.03571	0.14706	0.10526
151	0.35	0.29412	0.08333	0.2	0.05263	0.07692	0	0	0	0.07143	0.14706	0.26316
154	0.225	0.11765	0	0.075	0.10526	0.61538	0.05	0	0	0.07143	0.02941	0.02632
157	0.075	0	0	0.125	0.23684	0.07692	0	0	0	0.10714	0	0.02632
160	0.025	0	0	0.125	0.10526	0.23077	0	0	0	0.10714	0.20588	0.02632
163	0	0	0	0	0.05263	0	0	0.17647	0.04545	0.07143	0	0.13158
166	0	0	0	0.1	0	0	0	0	0	0.10714	0	0.15789
169	0	0	0	0.025	0.02632	0	0	0	0	0	0.02941	0.13158

Locus AAT64	EFL	NVA	SBI	CLA	BTX	KFL	TMX	CMT	WTX	STX	NNC	GTX
101	0	0	0	0.05263	0	0	0	0	0	0	0	0
104	0	0.0625	0	0	0.02941	0	0	0	0	0	0	0
107	0	0	0	0	0.05882	0	0	0	0	0	0	0
110	0	0	0	0	0	0	0	0	0	0	0	0.02632
56	0	0	0.08333	0	0	0	0.025	0	0	0	0	0
68	0.2	0	0	0	0	0	0	0	0	0	0	0
71	0	0	0	0.05263	0	0	0	0	0	0.10714	0	0
74	0.2	0.28125	0.25	0.10526	0.14706	0.22727	0	0.33333	0.13636	0.42857	0	0.36842
77	0.35	0.03125	0.5	0.05263	0.02941	0	0.6	0.33333	0.86364	0.07143	0.44444	0.02632
80	0.05	0.21875	0	0.23684	0.14706	0	0	0.16667	0	0.07143	0.27778	0.18421
83	0.025	0	0.08333	0.21053	0.11765	0	0	0	0	0.28571	0	0.10526
86	0.075	0	0.08333	0.15789	0.23529	0.22727	0.275	0.03333	0	0.03571	0.11111	0.18421
89	0.025	0.09375	0	0.05263	0.02941	0.09091	0.025	0	0	0	0.05556	0.10526
92	0.05	0.0625	0	0.07895	0.17647	0	0.075	0.1	0	0	0	0
95	0	0	0	0	0	0.27273	0	0.03333	0	0	0.11111	0
98	0.025	0.25	0	0	0.02941	0.18182	0	0	0	0	0	0