Southern Illinois University Carbondale

OpenSIUC

8-1-2019

Population Genetics of Rice Rats (Oryzomys palustris) at the Northern Edge of the Species Range

Phillip Conrad Williams
Southern Illinois University Carbondale, radwil@gmail.com

Follow this and additional works at: https://opensiuc.lib.siu.edu/theses

POPULATION GENETICS OF RICE RATS (*ORYZOMYS PALUSTRIS*) AT THE NORTHERN EDGE OF THE SPECIES RANGE

by

Conrad Williams

B.S. University of Texas at Austin, 2011

A Thesis
Submitted in Partial Fulfillment of the Requirements for the
Master of Science Degree

Department of Zoology
in the Graduate School
Southern Illinois University Carbondale
August 2019

THESIS APPROVAL

POPULATION GENETICS OF RICE RATS (*ORYZOMYS PALUSTRIS*) AT THE NORTHERN EDGE OF THE SPECIES RANGE

by

Conrad Williams

A Thesis Submitted in Partial

Fulfillment of the Requirements

For the Degree of

Master of Science in the Field of Zoology

Approved by:

Dr. Kamal M. Ibrahim, Chair

Dr. Edward Heist

Dr. Carey Krajewski

Graduate School
Southern Illinois University Carbondale
July 3, 2019

AN ABSTRACT OF THE THESIS OF

Conrad Williams, for the Master of Science degree in ZOOLOGY, presented on July 3, 2019, at Southern Illinois University Carbondale.

TITLE: POPULATION GENETICS OF RICE RATS (*ORYZOMYS PALUSTRIS*) AT THE NORTHERN EDGE OF THE SPECIES RANGE

MAJOR PROFESSOR: Dr. Kamal M. Ibrahim

The marsh rice rat (*Oryzomys sp.*) is a semiaquatic rodent native to wetlands in the southeastern United States. The northwestern-most part of the rice rat's range extends to Illinois where rice rats are found in wetlands across the southern part of the state. Recent studies have shown that rice rats in the United States can be divided into two species: *O. palustris* and *O. texensis*, but the taxonomic status of rice rats in Southern Illinois is unclear. To resolve this, I sequenced cytochrome-b and the control region, two regions of mitochondrial DNA, for 16 rice rats and constructed a phylogeny using these new sequences and previously obtained *O. palustris* and *O. texensis* sequences. In contrast to previous morphological assessments, I found that rice rats in Southern Illinois should be classified as *O. texensis*. This would extend the range of *O. texensis* north and west from its current extent. Further investigation using nuclear loci will be needed to confirm this classification.

Recent studies in Illinois have shown that rice rats metapopulation dynamics are dependent on the hydrology of the wetland that they occupy, but little is known about how these differences affect the genetic connectivity of rice rats. Specifically, a study on a floodplain site with dynamic changes in water levels and a reclaimed mining with a more stable hydrology showed different patterns of colonization and extinction of patches within each site. To investigate how the genetic structure of rice rats at these sites is affected by this dynamic, I genotyped 148 rice rats at 7 polymorphic microsatellite loci. I did not detect significant genetic differentiation between rice rat populations at the floodplains site and mining site nor between patches within these sites. Assignment testing with STRUCTURE did result in most, but not all, individuals being assigned to groups corresponding to the population from which were

i

sampled. Tests for population bottlenecks at both sites revealed that that one of the two study sites, MMRWFS, may have experienced a bottleneck event in the recent past. This would corroborate a previous finding based on demographic data.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE   PAGE

CHAPTER 1:

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

Biological taxa exhibit varying degrees of genetic divergence on a continuum from deep divisions between species to minor divisions between locally distributed populations of conspecifics. While much can be inferred about the evolutionary history of a given taxon by examining genetic differences at one end of this continuum or the other, a more complete understanding is gained by studying both. Phylogeography concerns the relationship between historical geographic phenomena and the spatial distribution of taxa (Avise 2001; Hickerson *et al.* 2010), thus examining relationships between more distantly related taxa. Phylogeography also encompasses the study of conspecific populations when there is historical-geographic component explaining the genetic structure of those populations. On finer spatio-temporal scales, the study of gene flow between conspecific populations on the modern landscape is typically considered the realm of population genetics.

Metapopulations are groups of populations of the same species that are distributed across a landscape and between which there is limited gene flow. Within each population or subpopulation, depending on what level of genetic divergence is being examined, individuals are assumed to be randomly mating. Metapopulations have long been considered in the context of population genetics (Wright 1931; Slatkin 1977), where the focus is on geneflow between subpopulations, as well as in a demographic context with focus on the dynamics of subpopulation extinction and recolonization (Levins 1969). Early models of metapopulations did not consider explicit geospatial locations of populations or the effects of the habitat matrix between populations. To account for these factors, researchers turned to landscape ecology and, by extension, landscape genetics.

Rice rats (*Oryzomys*) are sigmodontine rodents native to North and South America. They are part of tribe Oryzomyini, a taxon that arose and diversified in the Amazonian bioregion region of South America

(Maestri *et al.* 2018). North American rice rats are semiaquatic, wetland-living rodents that inhabits

coastal marshes and inland wetlands in the southeastern United States (Wolfe 1982). Rice rats in

Southern Illinois, the northwestern-most edge of their range in North America, form metapopulations in

wetlands with stable hydrological regimes or bordering large river floodplains where periodic inundation

occurs (Eubanks et al. 2011). In the present study, I examine the genetic structure of two rice rat

metapopulations in Southern Illinois using 7 polymorphic microsatellite loci. Additionally, I examine the

geographic origin of rice rats in Southern Illinois using mitochondrial DNA sequences.

RICE RATS IN SOUTHERN ILLINOIS

   Oryzomys palustris is a semiaquatic cricetid rodent that occurs primarily in wetlands along the

Atlantic and gulf coasts, but is also found throughout the inland southeastern United States at lower

densities in freshwater wetlands (Wolfe 1982). Recent studies of molecular divergence within *O.

palustris* indicate two distinct clades: an eastern clade consisting of *O. palustris*, and a western clade

consisting of *O. palustris texensis* (Hanson *et al.* 2010; Indorf 2010). Following the recommendation of

Hanson *et al.* 2010, I henceforth refer to the western clade as *O. texensis* and the eastern clade as *O.

palustris*, though reproductive isolation between the two clades has not been confirmed to date.

   Within the eastern clade, there are 6 typically recognized subspecies of *O. palustris* (Wolfe 1982). *O.

p. natator* and *O. p. coloratus* from central and southern peninsular Florida, respectively. *O. p. sanibeli*

and *O. p. planirostris* occupy Sanibel Island and Pine Island off the Florida coast. *O. p. argentatus*, the

silver rice rat, is a subspecies from the Lower Florida Keys. O. p. argentatus is sometimes referred to in

the literature as a separate species: *O. argentatus*. *O. p. texensis* occupies a range from Louisiana west

to eastern Texas and north to eastern Oklahoma. Finally, *O. p. palustris* occupies the bulk of the species

range from the panhandle of Florida, north to New Jersey, east to Mississippi, and north to Kentucky.

Morphologically, these subspecies are described as differing in size and pelage. Indorf and Gaines (2013)

found that the Pine Island rice rat likely wasn't genetically distinct enough to be considered a subspecies, but the Sanibel Island Rice rat does appear genetically distinct.

The phylogeographic history and distribution of rice rats in the southeastern United States was examined by Indorf (2010) using mitochondrial cytochrome b and control region sequences. Indorf's study indicates the divergence between *O. palustris* and *O. texensis* occurred during the early Pleistocene, 0.75 to 2.5 mya, likely as a result of separate glacial refugia (Indorf 2010). *O. palustris* appears to have expanded from an eastern glacial refugia to occupy a range including the United States east coast from Florida to New Jersey, and Alabama, Tennessee, and Kentucky. *O. texensis* expanded from a more western glacial refugia to occupy the Texas and Louisiana Gulf coasts as well as Mississippi, Arkansas, and eastern Oklahoma. However, this study did not include rats from Southern Illinois, leaving the geographic origin of these populations uncertain.

The first records of rice rats in Illinois come from Alexander County at the extreme southwestern edge of the state (Cory 1912; Necker and Hatfield 1951). The only extensive survey of the rice rat's range in Illinois was conducted in 1987 and found that the species was limited to the southern counties: Alexander, Franklin, Hamilton, Jackson, Johnson, Pope , Massac, Saline, White, and Williamson (Hofman *et al.* 1990). Necker and Hatfield (1951) list the rice rats present in Illinois as *Oryzomys palustris palustris*. Current range maps also show rice rats in Southern Illinois as belonging to *O. palustris palustris* (Wolfe 1982; Humphrey and Setzer 1989; See Figure 1 of Hanson *et al.* 2010), However, the first two studies did not use molecular data to confirm the subspecies status of rice rats in Southern Illinois, and the last (Hanson *et al.* 2010) did not include samples from Southern Illinois.

In Southern Illinois, rice rats occupy patchy habitat in inland wetlands with stable hydrological regimes and wetlands bordering large river floodplains where periodic inundation occurs. Rice rats often found in emergent wetlands and in ditches. It is associated with presence of upland grass cover,

herbaceous cover within the habitat, and dense, emergent vegetation such as cattails (*Typha sp.*),

sawgrass (*Cladium jamaicense*), and common reed (*Phragmites australis*) (Eubanks *et al.* 2011). Two

metapopulations in Southern Illinois have been extensively studied: one at a reclaimed mining site that

has been converted to wetland habitat in Jackson County (Burning Star 5; BS5), and another in a

floodplain of the Mississippi River in Alexander County (Middle Mississippi River Wetland Field Station;

MMRWFS). The floodplain metapopulation appears to exhibit a habitat-tracking dynamic where patch

occupancy is driven by changes in patch quality due to flooding. In contrast, the reclaimed mining site

metapopulation, which has a stable hydrological regime, better approximates a classic ecological

metapopulation (Van der Merwe *et al.* 2016)in which populations occasionally go locally extinct only

later to be recolonized by individuals from nearby patches.

 While *O. palustris* populations in other locations have shown little change in population size during long

periods (≥5 months) of inundation (Abuzeineh *et al.* 2007), populations in Southern Illinois floodplains

appear to be reduced by long periods of deep inundation (>2 meters) and to recover after inundation

(Van der Merwe *et al.* 2016). *O. palustris* individuals in Southern Illinois are highly vagile compared to

other rodents of similar size. A study of landscape permeability and movement showed an average daily

movement distance of 32.9 m, with a maximum recorded movement of >300 m in a single night (Cooney

2013). Movements of > 300 m have also been recorded in other areas of the species range (Negus *et al.*

1961), including the crossing of >300 m of open water (Forys and Dueser 1993). How this combination

of high vagility and patch colonization dynamic affect the genetic structure of populations remains an

open question. Genetic interconnectedness can be achieved with much lower migration rates than

demographic interconnectedness (Allendorf *et al.* 2012), and, therefore, it is important to examine the

two processes independently.

PHYLOGEOGRAPHY

   Phylogeography is a sub-discipline of biogeography that examines the "principles and processes governing the geographical distributions of genealogical lineages, especially those at the intraspecific level" (Avise 1998, pg 371). As traditionally practiced, phylogeography uses molecular markers to examine the causal relationship between historical geographic phenomena and the spatial distribution of taxa (Avise 2001; Hickerson *et al.* 2010). Mitochondrial DNA (mtDNA) has unique properties that make it ideal for examining phylogeographic relationships both within and between species. Mitochondrial DNA is haploid, maternally inherited, rapidly evolving, and non-recombining (Avise *et al.* 1987). These traits make mtDNA ideal for examining closely related populations and species (Avise *et al.* 1987).

   Currently, all literature seems to classify rice rats in Southern Illinois as belonging to *Oryzomys palustris* as opposed to the more western *O. texensis* (Wolfe 1982; Humphrey and Setzer 1989; Indorf 2010)*.* However, this relationship has never been examined using molecular data. A phylogenetic analysis based on mitochondrial DNA could help clarify which of these two clades rice rats in Southern Illinois belong to and add to our understanding of the phylogeography of North American *Oryzomys*.

POPULATION GENETICS

   Wright (1931) proposed measures of genetic differentiation, known as inbreeding coefficients, for various levels of hierarchical population substructure. Assuming genetic markers are selectively neutral, differentiation between demes at each of these levels is due to a balance between genetic drift, which is inversely proportional to the effective population sizes of demes, and gene flow among demes. The three most commonly recognized levels of structure are individual organisms (I), subpopulations (S), and the total population (T), although inbreeding coefficients can be extended to any number of nested levels of population structure.

F_{st} is a measure of genetic differentiation between demes or populations and is defined as $F_{st} = 1-$ (H_S/H_T) where H_S is the expected heterozygosity within a deme or subpopulation assuming Hardy-Weinberg equilibrium, and H_T is the expected heterozygosity computed from the pooled allele frequencies of the demes or populations. Deviation from Hardy-Weinberg proportions and, thus, higher F_{st} values can be caused by the Wahlund effect, an artifact of sampling that occurs when a sample of individuals is unknowingly drawn from two or more subdivided populations and treated as a single population. If a sample is drawn from a population where genetic substructure is present (e.g. treating two distinct populations as a single population), the sample population will appear to have excess homozygosity. Deviations from Hardy-Weinberg proportions can also be due to subdivided populations, nonrandom mating (Allendorf et $al.$ 2012). Holsinger and Weir (2009) reviewed F_{ST} and its applications.

Two other commonly used F statistics are F_{IS} and F_{IT}. F_{IS} is a measure of genetic differentiation due to nonrandom mating within a population (or a sub-population) and is defined as $F_{IS} = 1- (H_O/H_S)$. H_O is the mean observed heterozygosity within individuals and H_S is the expected heterozygosity based on allele frequencies within each deme. F_{IT} is defined as $F_{IT} = 1- (H_O/H_T)$ where H_O is defined as above and H_T is the expected heterozygosity based on allele frequencies in the total population (i.e. all demes pooled).

MICROSATELLITES

Microsatellites are DNA sequences consisting of short, tandem repeats of bases. The repeated unit usually consists of 1-3 nucleotides, but longer repeat units can occur. Microsatellite alleles vary in sequence length due to replication slippage, thus adding or subtracting one or more repeat units whenever a mutation occurs; this allows microsatellite evolution to be modeled as a stepwise mutation process (Ellegren 2004). Microsatellites have found wide use as codominant markers in population genetics. While microsatellites are prone to some degree of homoplasy due to alleles that are identically sized but not identical by descent, it is not likely to be severe enough to interfere significantly with many types of population genetic analysis (Estoup et $al.$ 2002).

One of the advantages of polymorphic microsatellite loci is their utility in detecting recent population bottlenecks that would otherwise be missed by demographic studies of population size (Allendorf *et al.* 2012). Population bottlenecks cause a reduction in the number of alleles relative to the level of heterozygosity at polymorphic loci. This reduction is characterized by the relative absence of low frequency alleles compared to that expected had a bottleneck not occurred (Luikart *et al.* 1998). The ability to detect a bottleneck depends on the severity of the bottleneck, time since the reduction in population size, and mutation model of the locus being examined. More severe, recent bottlenecks are easier to detect and it typically requires more loci to detect a bottleneck using loci following the stepwise mutation model than with loci following the infinite allele model (Cornuet and Luikart 1996).

Rice rats in Southern Illinois, though highly vagile (Cooney 2013), might provide an opportunity to examine the population genetic structure of a small, semi-aquatic rodent at the periphery of its range. With this in mind, I have chosen to examine this at two sites in Southern Illinois: Burning Star 5 (BS5), a reclaimed mining site in Jackson County and the Middle Mississippi River Wetland Field Station (MMRWFS), located in a floodplain of the Mississippi River in Alexander County. The differing hydrological regime between the two sites (inundation occurs much more frequently at MMRWFS) further provides the opportunity to examine the effects of flooding on genetic connectivity.

CHAPTER 2:

STUDY SITES AND SAMPLE ACQUISITION

STUDY SITES

   The rice rat samples used in the population genetics study were collected from two wetland

complexes in Southern Illinois. The first site is a reclaimed mining site at the Illinois Department of

Natural Resources CONSOL Energy - Burning Star 5 Wildlife area (hereafter referred to as Burning Star 5

or BS5) in Jackson County (37° 52' 32.95" N; 89° 12' 30.47" W). Most mining at Burning Star 5 occurred

after 1978 when reclamation of at least part of the mined areas was required by federal law. Burning

Star 5 is located in the Big Muddy River watershed, and covers approximately 3,200 ha, including

roughly 1,200 ha of upland wetland habitats and 640 ha of unmined forested areas (Figure 1, Figure 2).

   The second site is located at the Middle Mississippi River Wetland Field Station (hereafter MMRWFS)

in Alexander County in Illinois. Managed by Southern Illinois University, MMRWFS covers approximately

560 ha and is located on the east side of the Mississippi River approximately 4 km East-South-East of

Cape Girardeau, Missouri (37° 17' 2.08" N; 89° 28' 6.37" W). The area consists of various managed

natural wetlands within the Cape Bend State Fish and Wildlife Area and the Shawnee National Forest

(Figure 1, Figure 3). A primary difference between the conditions at MMRWFS and Burning Star 5 is that

MMRWFS experiences periodic inundation due to occasional flooding of the Mississippi River and

proximity of the water table to the surface. Water depths in wetlands at BS5 vary with season and

precipitation, but the area does not become totally inundated.

   In addition to the above 2 sites, a small number of rice rats used in the phylogeography study were

collected from a site approximately 5 km west of Harrisburg, Illinois in Saline County (37° 44' 31.00" N;

88° 36' 24.10" W). This site is another reclaimed mining site and is surrounded by residential areas,

cropland, and abandoned cropland (Hofmann and Gardner 1992). Due to a very small number of

samples from this site, it was not included in the population genetic analyses.

SAMPLE ACQUISITION

Rice rat tissue samples used in the current studies were obtained by Dr. Jorista van der Merwe during 2011-2013 (van der Merwe 2014, van der Merwe et al. 2016). Rice rats were captured using Sherman live traps (8 cm x 9 cm x 23 cm, H.B. Sherman Traps Inc., Tallahassee, Florida, USA). Traps were deployed in transects of 30-100 traps adjacent to 9 wetland patches at MMRWFS and 14 wetland patches at BS5. Each trap was baited with birdseed. Trapping sessions took place 4 times per year between March 2011 and November 2013: spring (mid-March to mid-May), early summer (mid-May to mid-July), late summer (mid-July to mid-September), and fall (mid-September to mid-November). During each session, traps were deployed for 3-4 consecutive nights at each site.

In total, 785 and 193 individual rice rats were captured at MMRWFS and BS5, respectively. One toe on either of the forepaws of each rat was clipped at the first joint to mark the animal and provide a tissue sample for DNA extraction. Samples were collected in accordance with guidelines approved by the SIUC Institutional Animal Care and Use Committee (IACUC Protocol 10-009; Sikes *et al.* 2007). Toe clips were preserved in 95% ethanol and frozen until shortly before DNA extraction.

CHAPTER 3:

THE PHYLOGEOGRAPHIC ORIGINS OF RICE RATS IN SOUTHERN ILLINOIS

INTRODUCTION

Recent studies of molecular divergence among *Oryzomys palustris* populations recommend dividing them into two separate species: an eastern clade called *Oryzomys palustris* and a western clade called of *Oryzomys texensis* (Hanson *et al.* 2010; Indorf 2010). The phylogeographic history and distribution of rice rats in the southeastern United States was examined by Indorf (2010) using two mitochondrial regions (cytochrome b and the control region). Indorf's study indicates that the divergence between *O. palustris* and *O. texensis* occurred during the early Pleistocene, likely as a result of isolation in separate glacial refugia prior to northward range expansion as ice sheets receded and suitable habitat became available. However, this study did not include rice rats from Southern Illinois, leaving the origin of these populations uncertain.

In the absence of genetic data, current knowledge of the history of rice rats in Southern Illinois relies on fossils, species records, and morphological comparisons. Fossil remains indicate that rice rats were present in the southern US during the Pleistocene (Wolfe 1982). Rice rat remains from most archeological sites at the northern extent of their range date to ~1000 A.D. and suggest that the rats may have been a commensal pest of native American maize crops in the Midwest (Richards 1979). Living rice rats were first recorded in freshwater swamps of the eastern US in 1837 (Harlan 1837) and first recorded in Southern Illinois by Howell (1910). These and other records were incorporated into a range map by Goldman (1918), who assigned rice rats in Southern Illinois to *O. palustris palustris* (and also, notably, recognized *O. palustris texensis* as a western subspecies on the same map). Goldman's classification was used in a dichotomous key for mammals in Illinois Necker and Hatfield (1951) and is referenced in Humphrey and Setzer's (1989) attempt to classify rice rats into subspecies based on morphological characteristics (Necker and Hatfield 1951; Humphrey and Setzer 1989).

The species level differences recognized by Indorf (2010) stand in stark contrast to Humphrey and Setzer's (1989) study based on morphology alone, which suggested *O. palustris* and *O. texensis* be classified as a single subspecies: *O. palustris palustris*. This incongruence between molecular and morphological results calls into question the taxonomic status of rice rats in Southern Illinois.

In this study, I use new DNA sequence data from two mitochondrial regions (cytochrome-b and the control region), along with existing sequence data from GenBank, to investigate the phylogeography of rice rats in Jackson, Alexander, and Saline counties of Southern Illinois. I find that these populationas represent a northward expansion of *O. texensis*, not O. *palustris palustris* as indicated by current range maps and previous morphological assessments (Mclaughlin and Robertson 1951; Hanson *et al.* 2010; Indorf 2010; Indorf and Gaines 2013).

METHODS

DNA extraction, amplification, and sequencing: I extracted DNA from 16 rice rat toe clips (6 from BS5, 5 from MMRWFS, and 5 from the Harrisburg site; see Chapter 2 for site descriptions) using Qiagen DNeasy tissue kits (Qiagen Inc.) following the manufacturer's instructions for tissue. I then amplified two mitochondrial regions, cytochrome-b (Cytb) and the control region (CR), following procedures modified from Indorf (2010). Polymerase chain reaction (PCR) primers for CR were 2340-5 (Mendez-Harclerode *et al.* 2005) and Ory5' (Indorf and Gaines 2013). For Cytb, I initially used MVZ05 (Smith and Patton 1993) and CB40 (Hanson and Bradley 2008). These initial attempts failed. However, by using MVZ05 and the reverse complement of CB40, I was able to obtain the expected PCR product. PCR primer sequences are shown in Table 1. For both loci, PCR reactions totaled 10 microliters and consisted of 1.4 μl of each primer (final concentration 1.4 μM), 0.5 μl bovine serum albumin (BSA; 20mg/ml), 0.7 μl water, 1 μl DNA extract, and 5 μl Dreamtaq Master Mix (2X; Thermo Scientific). For CR, the thermocyler program consisted of the following: a 5-minute initial denaturing at 95°C, 35 cycles of 95°C denaturing (1 minute), 49°C annealing (45 seconds), and 72°C extension (90 seconds), and an 8-minute final extension. A similar

program was used for Cytb: a 5-minute initial denaturing at 95°C, 35 loops of 95°C denaturing (1

minute), 54°C annealing (45 seconds), and 72°C extension (90 seconds), and an 8-minute final extension.

I visualized the PCR products on agarose gel stained with ethidium bromide prior to enzymatic

cleanup of successful reactions with ExoSAP-IT (Affymetrix). I used the BigDye Terminator v3.1 Cycle

Sequencing Kit (Applied Biosystems) for cycle sequencing reactions. Cycle sequencing products were

purified using Illustra Sephadex spin columns (GE Healthcare) and dried for 30 minutes in a vacuum

centrifuge before being resuspended in 10 µl Hi-Di formamide (Applied Biosystems) and resolved on an

ABI 3130XL genetic analyzer (Applied Biosystems). For both PCR products, I ran separate cycle

sequencing reactions for both forward and reverse primers, thus obtaining overlapping, reverse

complementary sequences for each locus that were later trimmed and assembled in Geneious 10.1.3

(Kearse *et al.* 2012).

Data Analysis: In addition to my 16 Cytb and CR sequences, I downloaded 3 sets of data from

GenBank's PopSet database (PopSet IDS: 403390915, 403390649, and 288223063). These PopSets

contained rice rat Cytb and CR sequences from multiple sites in the southeastern United States and

Mexico that were used in phylogenetic and phylogeographic studies by Hanson *et al.* (2010), Indorf

(2010), and Indorf and Gaines (2013). From these sets, I filtered out sequences from species other than

O. palustris and *O. texensis* (retaining some sequences from *O. couesi* as an outgroup for later analysis).

Individuals, that did not have sequences for both loci were identified by voucher number and excluded.

After this initial filtering step, I retained 235 sequences for each locus from the PopSets in addition to

my 16 sequences for each locus. I aligned my sequences and the sequences derived from the PopSets

for each region separately using the MUSCLE plugin in Geneious 10.1.3 (Edger R 2004; Kearse *et al.*

2012). I then concatenated the two alignments based on voucher number and my own sample IDs.

Finally, I eliminated duplicate haplotypes from the concatenated matrix leaving 116 concatenated Cytb

and CR sequences. This alignment was used for all subsequent analyses.

Selecting appropriate models of molecular evolution is important to obtaining accurate estimates of phylogenetic relationships. I ran the above alignment through PartitionFinder 2.1.1, a program that aids in selecting the best-fitting partitioning scheme and models of molecular evolution for a given alignment (Guindon *et al.* 2010; Lanfear *et al.* 2016). I limited the models tested to include only models that could be implemented in MrBayes and allowed four putative partitions: the control region and each codon position of Cytb. I used BIC as the model selection criterion. The best partitioning scheme used all four putative partitions. The best partitioning scheme was used in both the MrBayes and RAxML analyses while the best models of molecular evolution for each partition were only used in the MrBayes analysis.

I used two tree inference methods to construct phylogenetic trees from the above alignment. For maximum likelihood analysis, I used the rapid bootstrapping method implemented in RAxML 8.2.9 (Stamatakis *et al.* 2008; Stamatakis 2014) on CIPRES science gateway v.3.3 (Miller *et al.* 2010). For this analysis, I partitioned the dataset into the control region and each codon position of Cytb and used RAxML's GAMMACAT model of molecular evolution for all partitions under default settings. I performed 1000 bootstrap replicates to obtain support values for each node.

For the Bayesian analysis, I used MrBayes v3.2.6 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003; Altekar *et al.* 2004) on CIPRES. I used both the partitioning scheme and models of molecular evolution selected by PartitionFinder. Settings for the Markov Chain Monte Carlo (MCMC) were as follows: 2 independent runs, 4 chains, 1 million generations with trees sampled every 1000 generations. I used the potential scale reduction factor (PSRF) reported by MrBayes to assess chain convergence (Gelman and Rubin 1992). This statistic is expected to approach 1 as chains converge. The first 25% of sampled trees were discarded as burn-in. The remaining sampled trees (a total of 3002, 1501 from each independent run) were summarized as a 50% majority rule consensus with clade credibility values included for the bipartitions using the "sumt" command.

My analysis uncovered an unexpected tree topology given our current understanding of the distribution of *Oryzomys* in the United States. One way to test whether the topologies I recovered are better explanations for my data than the topology suggested by current range maps (Humphrey and Setzer 1989; Hanson *et al.* 2010) is to compare the results of my unconstrained analysis to those of an analysis constraining rice rats from Illinois to form a clade with *Oryzomys palustris*. The SOWH test is a likelihood-based parametric bootstrapping method that allows comparison of specific alternative topologies. I used SOWHAT (Church *et al.* 2015), a program that implements the SOWH test, to test for the monophyly of a clade consisting of rice rats from Illinois and *Oryzomys palustris*. I provided SOWHAT with a constraint tree assigning all samples from Illinois and *Oryzomys palustris* to one clade and all samples from *Oryzomys texensis* to a sister clade along with the alignment used in all other analyses (described above). I also provided SOWHAT with the same alignment partitions described in the above RAxML analysis. SOWHAT uses Seq-Gen v1.3.3 (Rambaut and Grass 1997) to simulate 1000 alignments and RAxML v8.1.20 (Stamatakis *et al.* 2008; Stamatakis 2014) to infer topologies for each of those data sets under both constrained and unconstrained maximum likelihood analysis. The differences in likelihood between the constrained and unconstrained trees inferred from simulated data form a null distribution that is used to test the significance of the difference in likelihoods between constrained and unconstrained analyses from my alignment.

RESULTS

The 16 newly obtained Cytb sequences were all 1,123 bp (Appendix A). The control region from 15 of the newly-sequenced individuals ranged 938 bp to 960 bp (Appendix B). The reverse sequencing reaction for a single individual (HAR1) failed. The control region sequence for this individual was truncated at 887 bp due to low confidence low confidence in base calls at the 3'-end of the forward sequence. I found three different combined Cytb and CR haplotypes in Southern Illinois. A haplotype

shared by 11 of individuals was indistinguishable from the haplotype of an individual sampled from Shelby County in Tennessee. The other two haplotypes are unique to Southern Illinois.

PartitionFinder showed the best partitioning scheme was placing the control region and each codon position of Cytb into separate partitions (4 total partitions). The best model for the control region was GTR+I+G. HKY was the best model for the first codon position of Cytb, GTR+G for the second, and K80+I for the third.

The maximum likelihood analysis in RAxML and the Bayesian analysis in MrBayes both produced phylogenetic trees with strong support for 2 major clades: one consisting entirely of *O. palustris* (*O. palustris palustris*, *Oryzomys palustris argentatus*, and *Oryzomys palustris sanibel*) and the second consisting of *O. texensis* and my samples from southern Illinois (Figures 4, 5). The difference in likelihood between the unconstrained analysis and one constraining rice rats in Southern Illinois to *O. palustris* was 95.0357; the SOWH test rejected the hypothesis of Illinois rice rats forming a clade with *Oryzomys palustris* ($p < 0.001$, 95% confidence interval: 0 - 0.00368).

DISCUSSION

The two major clades recovered in both phylogenetic analyses correspond to clades A and B in Hanson *et al.* (2010), one of the studies from which many of the GenBank sequences I used were derived. All Southern Illinois samples were assigned with high support to a clade with *Oryzomys texensis* (Hanson *et al.'s* clade B) and not to *O. palustris palustris* as current range maps and previous morphological assessments predicted (Mclaughlin and Robertson 1951; Hanson *et al.* 2010; Indorf 2010; Indorf and Gaines 2013). The earliest study of rice rat distribution in Southern Illinois (Cory 1912), as well as Necker and Hatfield's (1951) account, refers to rice rats in Southern Illinois as *O. palustris palustris*; these appear to be the origin of the current subspecies designations on range maps.

All sequences included in this study are mitochondrial. This raises the possibility that the results are due to mitochondrial introgression, implying that the Southern Illinois rice rats are actually *O. palustris palustris* but harbor mitochondrial DNA introgressed from *O. texensis*. Introgression at secondary contact zones between recently diverged lineages is common (Harrison and Larson 2014). Other small rodent species have shown asymmetric mitochondrial introgression within and even beyond contact zones between divergent lineages (Jaarola *et al.* 2011; Mathias *et al.* 2012). However, based on the data that I have gathered, I cannot conclude this has occurred. I only found haplotypes assignable to *O. texensis* in the 16 individuals from which I obtained new Cytochrome-b and Control Region sequences. Hanson et al. (2010) found a ~6% sequence divergence in Cytb between *Oryzomys texensis* and *Oryzomys palustris* which is comparable to interspecific variation between other mammal species and at least suggests a reproductive barrier between those two lineages (Bradley and Baker 2001; Hanson *et al.* 2010). Therefore, my best assumption is that rice rats in Southern Illinois belong to *O. texensis*.

Another consideration to bear in mind is that the sampling locations in this study were limited to sites previously used in the 2013 meta-population dynamics study by van der Merwe (2014). Broader sampling could have revealed additional sequence variability that impacted the phylogenetic analyses. This highlights the need for broader sampling within the region and combined mitochondrial and nuclear loci to determine conclusively the origins of rice rats in Southern Illinois.

POPULATION GENETICS OF RICE RATS IN SOUTHERN ILLINOIS

INTRODUCTION

Marsh rice rats (*Oryzomys sp.*) are present in several wetland complexes across Southern Illinois. The most recent survey of the distribution of rice rats in Southern Illinois, based on sampling at 49 sites, found that the presence of rice rats in wetlands was associated with upland grass cover, percent herbaceous cover, percent visual obstruction 0-0.5m from the ground (Eubanks 2009). The patchy nature of wetlands in Southern Illinois combined with the high vagility of rice rats (Cooney 2013) suggest that the rats may exist as metapopulations.

The demographics and colonization/extinction dynamics of rice rat metapopulations at two wetland sites were studied in depth by van der Merwe (2014). The Middle Mississippi River Wetland Field Station (MMRWFS) in Alexander County, Illinois experiences periodic inundation due to occasional flooding of the Mississippi River and a shallow water table. In contrast, the Illinois Department of Natural Resources CONSOL Energy - Burning Star 5 Wildlife area (BS5) in Jackson County has a more stable hydrological regime and is never completely inundated (see chapter 2 for full site descriptions). Van der Merwe (2014) predicted that tice rat metapopulations at these sites would differ in patterns of patch colonization and extinction depending on the hydrology of the habitats. However, this pattern failed to materialize, perhaps due to extremes of drought and flooding. Instead, rainfall was the best predictor of colonization rate. From 2011-2013, colonization rates and relative abundance were generally lower at BS5 than MMRWFS (Van der Merwe 2014). In 2011, severe flooding caused the extinction of many populations at MMRWFS with subsequent recolonization from refuges within the site, perhaps indicating a population bottleneck. Patch occupancy was more dependent on patch size than on the land cover type of the surrounding matrix habitat. Overall, MMRWFS appeared to exhibit a classic metapopulation dynamic while BS5 showed a habitat-tracking dynamic, in which suitable patches are

colonized as they become available and go extinct due to environmental changes (e.g. flooding or drought).

The same processes that drive differences in demographics and patch colonization between rice rat metapopulations at these sites might cause differences in genetic structure. For example, the rice rat populations at BS5 might be more homogenous than those at MMRWFS because a true metapopulation should theoretically harbor greater genetic diversity than the habitat-tracking dynamic at BS5. Understanding these effects could have implications for assessing what wetland sites within Illinois are better for conserving rice rat genetic diversity. To that end, I genotyped rice rats from multiple hypothesized populations within each wetland site at 7 polymorphic microsatellite loci. Using this dataset, I examined possible population bottleneck effects and genetic structure within and between sites.

METHODS

Study Sites, Sampling, and DNA extraction – The actual boundaries of rat populations at each site are unknown. Therefore, I subdivided the sites into hypothetical populations based on proximity of trap line locations (see chapter 2 for site descriptions and trapping procedures) and the number of samples that were available at each site (figure 4). I extracted DNA from 20-32 rice rat toe clips from each population using Quiagen DNeasy tissue kits (Quiagen Inc.) following the manufacturer's instructions for tissue, with the exception of reducing elution from 200 microliters to 100 microliters to increase DNA concentration and compensate for the small size of my tissue samples (Table 2). In total, I extracted DNA from 148 toe clips spanning a 3-year sampling period from 2011-2013.

Microsatellite Genotyping – I attempted to amplify 9 polymorphic microsatellite loci developed specifically for *O. palustris* (Wang *et al.* 2000). I was able to successfully amplify 7 of these from my samples after multiple rounds of optimization. The 5' end of the forward primer for each locus was

modified with a fluorescent label as indicated in table 3. The PCR program I used only differed among loci in annealing temperature (listed in table 3) and was as follows: 5 minutes 95 °C; 42 cycles of 45 seconds at 95 °C, 60 seconds annealing, and 45 seconds at 70 °C; 7 minutes at 70 °C. Loci with similar annealing temperatures were multiplexed. After amplification, PCR products were analyzed on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Alleles were scored in the Geneious 11.1.5 microsatellite plugin version 1.4.4 (Biomatters Ltd.) using automated allele binning and checked by hand for incorrect peak calls. For each sample, any locus that failed initial amplification was attempted 3-4 more times in singleplex. Any locus still failing at that point was regarded as missing data in subsequent analyses.

Analysis – I used Arlequin 3.5.2.2 (Excoffier *et al.* 2005) to examine genetic variation within and among the putative populations. For each putative population, I used Arlequin to calculate expected and observed heterozygosity, F statistics including pairwise F_{ST} (Wright 1931), and allelic richness. Arlequin was also used to perform analysis of molecular variance (AMOVA) (Weir and Cockerham 1984; Excoffier *et al.* 1992; Weir 1996). AMOVA examines the genetic distance between individuals and assigns variation in this distance to different levels of hierarchical population structure. In this case, MMRWFS and BS5 represent two groups with 3 putative populations within each (see Table 2).

Initial analysis revealed a significant excess of homozygosity at many loci within each putative population. Therefore, I used Micro-Checker (Van Oosterhout *et al.* 2006) to test for the presence of null (non-amplifying) alleles, large allele dropout, and stuttering (in vitro mutation during PCR) that can cause deviations from Hardy-Weinberg equilibrium (HWE). While Micro-checker can give estimates for adjusted allele frequencies and adjusted genotypes, I chose not to use these in subsequent analyses. Micro-checker bases its estimated allele frequencies on the assumption that the population is in HWE (Van Oosterhout *et al.* 2004). Since I cannot say with certainty that the populations in this study meet that assumption, it would be unsound to use the estimated frequencies. The results analyses using the original, unadjusted dataset are discussed in the context of results from Micro-Checker.

19

I used the program STRUCTURE 2.3.4 (Pritchard *et al.* 2000; Hubisz *et al.* 2009), a Bayesian clustering method, to perform assignment testing. Under its admixture model, STRUCTURE assigns individuals' ancestry to K populations or clusters while attempting to minimize deviations from HWE and linkage disequilibrium. I ran STRUCTURE with a burn-in of 10,000 iterations followed by 100,000 iterations of Markov Chain Monte Carlo (MCMC) for K = 1 to K = 8. Five independent replicates were performed for each K. I then used the program Structure Harvester (Earl and VonHoldt 2012) to calculate Evanno et al.'s ΔK (Evanno *et al.* 2005), a method for selecting the best estimate of the number of populations present in the dataset. I used CLUMPP to summarize independent runs at the K selected by the Evanno method and used DISTRUCT to visualize individual ancestry assignment (Rosenberg 2004; Jakobsson and Rosenberg 2007).

The results of the initial run of STRUCTURE recovered two distinct clusters at K = 2 that did not correspond to geographic distribution. To further explore the data, I then repeated the above procedure separately for both clusters using samples with greater than 90% of their ancestry assigned to a given cluster (110 individuals in the "Red" group from figure 6; 27 individuals in the "Green" group from figure 6). Further, I repeated the original procedure 7 more times, each time removing 1 of the 7 microsatellite loci to assess the influence of each locus on population assignment. In these analyses, I used the full-search algorithm in CLUMPP for K < 5. For K ≥ 5, I used the greedy search option with 10,000 independent starts. Finally, I ran one more STRUCTURE analysis on a 'strict' dataset consisting only of individuals with no missing data and with locus OryAAT21 removed. For this run using the 'strict' dataset, I ran STRUCTURE with a burn-in period 100,000 iterations followed by 1,000,000 iterations of MCMC for K = 1 to K = 8. Eight independent replicates were performed for each K to ensure convergence. I again used Structure Harvester, CLUMPP and DISTRUCT to summarize and visualize these runs as described for the first STRUCTURE analysis above.

To test for population bottlenecks at MMRWFS and BS5, I used the program BOTTLENECK (Piry *et al.* 1999). BOTTLENECK examines the relationship between reduction in heterozygosity and reduction in allelic richness at microsatellite loci. Theoretically, when a population bottleneck occurs, allelic richness is reduced faster than heterozygosity. This would result in heterozygosity excess in populations experienced recent bottlenecks. I ran the program separately for BS5 and MMRWFS and used the Wilcoxon's test statistic for assessing significance (this is the most appropriate test for datasets with fewer than 20 loci; Piry *et al.* 1999).

RESULTS

I successfully extracted DNA from 148 individual rice rats and genotyped them at 7 microsatellite loci. The 7 loci were all polymorphic in each of the 6 putative populations. The average number of alleles per locus across all populations ranged from 7.83 alleles for OryAAT26 to 13.50 alleles for OryAAT40 (Table 4). Across all loci, 117 different alleles were detected. Multiple loci in each population showed significant heterozygosity deficits (Table 5).

The Micro-Checker analysis detected between 3 and 4 loci with null alleles present in each population, but did not detect errors due to stuttering or large allele dropout. The estimated frequencies of null alleles using the Brookfield (1996) method ranged from 0.078 for OryAAT40 in population M3 to 0.314 for OryAAT10 in population B3 (Table 6). OryAAT03 was detected as having null alleles present in all 6 populations. These results indicate that it is likely that the heterozygosity deficit seen in all populations is due to technical error, specifically null alleles. While it is possible, under certain assumptions, to estimate new genotype frequencies from Micro-Checker's adjusted allele frequencies, I chose to analyze the original dataset as is and discuss those analyses in the context of Micro-Checker's results.

Pairwise F_{ST} values, measures of genetic differentiation, between pairs of populations ranged from -0.

016 to 0.031 (Table 7). The highest F_{ST} was between populations M3 and B3, and the lowest was

between B1 and B2, but F_{ST} between all populations at BS5 (B1, B2, and B3) was effectively 0. Most F_{ST}

estimates were significant under a permutation test performed in Arlequin, with the exception of all

pairs including population B1. F_{ST} estimates within MMRWFS were significant between M2 and M3, and

between M2 and M5. F_{ST} estimates were used to calculate the number of migrants per generation

between each pair of populations (Table 8) (Slatkin 1991). This calculation implicitly assumes equivalent

population sizes and bidirectional exchange of migrants, which might not be correct because these

populations likely differ in size (Van der Merwe 2014).

AMOVA (Table 9) indicated that 98.02% of genetic variation is attributable to variance within

populations (F_{ST} = 0.020, p < 0.000). Only 1.71% of genetic variation was attributed to differentiation

between BS5 and MMRWFS (F_{CT} = 0.017, p 0.112 ± 0.009) and only 0.27% of the variation was attributed

to differentiation among populations within sites (F_{SC} = 0.003, p = 0.463 ± 0.015). This indicates that

there is little genetic differentiation between BS5 and MMRWFS.

The STRUCTURE analysis yielded mixed results. Under the Evanno et al. (2005) criterion (Table 10),

K=2 was the best choice for the number of populations in the dataset. . However, the distribution of

ancestry assignment to these clusters (Figure 6) reveals no pattern corresponding to differentiation

specifically between MMRWFS and BS5.

I repeated the STRUCTURE analysis separately for individuals with >90% of their ancestry assigned to

"red" and individuals with >90% of their ancestry assigned to "green". For the "red" group (n=110), the

best number of clusters was K = 6 (Figure 7, A); for the "green" group (n=27), the best number of

clusters was K = 7 (Figure 7, B). "Green" individuals appear to have been separated out in the original

analysis due to a large amount of missing data. To that end, ancestry assignment is ambiguous for these

individuals under both K = 7 and K = 2. The "red" group shows a pattern that is better adheres to sampling location than the original analysis under both K = 6 and, especially, K = 2. Most individuals from MMRWFS are assigned mostly "orange" ancestry and most individuals from BS5 are assigned mostly "blue" ancestry. Mitochondrial haplotype does not reflect whether an individual was assigned "orange" or "blue" ancestry (table 11).

In order to examine whether results in the original analysis (Figure 6) were due to a specific locus, I reran the STRUCTURE analysis iteratively removing 1 of the 7 loci for each run (Figure 8). These reduced analyses gave similar results that resembled the unreduced analysis, with the exception of removing OryAAT21. Here, the best number of clusters was K =5. Under K=5, most individuals in BS5 were assigned 'green' ancestry while individuals at MMRWFS were assigned mostly 'blue' ancestry. Individuals assigned mostly 'yellow' ancestry have high amounts of missing data.

Because these analyses indicate the STRUCTURE results were sensitive missing data and removal of the OryAAT21 locus, I performed an analysis in which both OryAAT21 and any individuals with missing data were removed. These strict criteria reduced the data set to 84 individuals (44 from MMRWFS and 40 from BS5). In this analysis, the optimal K was K = 2. Most rice rats from MMRWFS were assigned 'orange' ancestry while most rice rats at BS5 were assigned 'blue' ancestry (Figure 9). These results are similar to the OryAAT21 removed dataset and do not correspond to mitochondrial haplotype (Table 12).

To examine how individual loci affected STRUCTURE assignment in this 'strict' dataset, I assigned rice rats with >50% 'blue' ancestry to one population (n = 39) and rice rats with >50% 'orange' ancestry to a second population (n = 45), and then examined allele frequencies for each locus using the Excel-Microsatellite-Toolkit version 3.1 (Park 2001). The assignment appears to be driven in part by two alleles of OryAAT03. Individuals homozygous for allele 124 were assigned to have mostly 'blue' ancestry, while individuals homozygous for allele 130 were all assigned orange ancestry, irrespective of which site the

individuals belonged to. This pattern could result from STRUCTURE trying to minimize deviations from HWE in the 'orange' and 'blue' groups by forcing all of the homozygotes for a given allele into the group where that allele is more common. For example, In the original 'strict' dataset, individuals that are homozygous for allele 124 appear at both MMRWFS and BS5 and there is an excess of homozygotes at both sites. STRUCTURE tries to minimize this deviation from HWE by assigning all of the 124 homozygotes to 'blue' thus raising frequency of allele 124 and the expected frequency of homozygotes. This results in less deviation from HWE in the 'orange' and 'blue' groups than in the 'strict' dataset.

BOTTLENECK gave mixed results depending on the mutation model assumed and further on the parameterization of the model. Microsatellite loci are best modeled by either the stepwise mutation model (SMM) or the two-phase model (TPM) (Piry *et al.* 1999). The only difference between these is that the SMM assumes mutations are always strictly one repeat unit while the TPM allows mutations of more than one repeat unit to occur at lower frequencies. TPM has been shown to be more appropriate for microsatellite data than the SMM (Piry *et al.* 1999). Under the SMM, BOTTLENECK failed to detect significant heterozygosity excess at neither MMRWFS (one-tailed Wilcoxon test: $p = 0.945$) or BS5 ($p = 0.988$). Under the TPM, BOTTLENECK gave different results depending on parameterization of the model. Under BOTTLENECK's default settings for the TPM (Probability of SMM in TPM = 70%, Variance for TPM = 30), the Wilcoxon test detected a heterozygosity excess at MMRWFS ($p = 0.003$) but not at BS5 ($p = 0.188$). This could indicate a past population bottleneck occurred at MMRWFS. However, under the parameterization recommended by Piry *et al.* (1999) (Probability of SMM in TPM = 95%, Variance for TPM = 12), the test was not significant for MMRWFS ($p = 0.234$) or BS5 ($p = 0.813$).

DISCUSSION

Overall, I detected no significant genetic differentiation between MMRWFS and BS5 (AMOVA F_{CT} = 0.017, $p = 0.112 \pm 0.009$). Nor is there genetic differentiation between putative populations within sites (AMOVA F_{SC} = 0.003, $p = 0.463 \pm 0.015$, and low pairwise F_{ST} estimates). This might seem surprising given

24

the distance between MMRWFS and BS5, but it is comparable to estimates of genetic differentiation

found in other studies of rice rats (Indorf 2010). Further, connectivity between the sites could be

explained by intermediate populations, such as those occupying roadside ditches (Eubanks *et al.* 2011).

The putative populations I assumed within each site are not sufficiently differentiated to warrant

considering them separate genetic populations, though they may behave as a metapopulation in a

demographic sense (Van der Merwe 2014; Van der Merwe *et al.* 2016). MMRWFS might exhibit some

genetic structure among populations as evidenced by significant pairwise F_{ST} estimates, but the

estimates themselves are mostly very low. This could be an effect of patch colonization / extinction

dynamics as shown by Slatkin (1977). Specifically, MMRWFS might fit what Slatkin called the propagule

pool model, in which patches are colonized by a small group of individuals from a single patch. If

flooding in 2011 nearly wiped out the rice rat population at MMRWFS, then it is possible that patches

were recolonized from a small refuge population on the Mississippi River levee. This scenario is

supported by rice rat captures only occurring near the levee in 2011, with captures at other MMRWFS

localities subsequently occurring in 2012 and 2013 (Van der Merwe 2014).

The potential detection of a population bottleneck at MMRWFS corroborates the prediction of van

der Merwe (2014). It is quite likely that the 2011 flood at MMRWFS severely reduced the rice rat

population and that the population was recovering during the sampling period in 2011-2013. No

bottleneck was detected at BS5, again consistent with demographic studies (Van der Merwe 2014).

However, the effect of null alleles on this assessment is not clear. The effect of null alleles on expected

heterozygosity as estimated by Bottleneck (i.e. $1-\Sigma p_i^2$; 1 minus the sum of the squared allele frequencies)

depends on the frequencies of null alleles relative to other alleles in the population, as well as whether a

null allele never amplified or amplified only in homozygotes (i.e. a partial null). Given that the signal for

the population bottleneck at MMRWFS is weak and that detection of that signal is highly dependent on

model choice and parameterization, the result may be unreliable in presence of null alleles.

The significant excess homozygosity across all populations in the study, but not for each locus within any given population, suggests the presence of unamplified alleles in many samples. This could be due to mutations in primer binding sequences flanking the microsatellite loci or to degradation of template DNA. Other studies using these microsatellite markers have also detected null alleles in many populations (Wang *et al.* 2005; Indorf 2010). Further, while I accounted for spatial structure when selecting samples, I did not account for temporal structure. If rice rat populations undergo patch colonization / extinction as has been hypothesized, samples from the same site at two different points in time are not necessarily drawn from the same population. This would introduce a Wahlund effect that could explain excess homozygosity. This, however, should affect all loci within a population; it is unlikely to be the case in this study because no populations showed deviations from HWE at all loci. Excess homozygosity may be partially attributable to inbreeding. However, given the aforementioned effects of null alleles and my sampling scheme, I cannot confirm whether or not either of these populations are actually inbreed.

The results of the initial STRUCTURE analysis were inconclusive. The Evanno *et al.* (2005) method for selecting the number of genetic clusters or populations present in the dataset yielded K=2 as the best fit. This seems reasonable given that there are two major sites within the study (MMRWFS and BS5), but the plot of ancestry assignment revealed that there is no pattern attributable to division between sites. Further analysis revealed that this partitioning was primarily between samples with large amounts of missing data and the rest of the dataset. Another set of STRUCTURE analyses revealed that results were sensitive to a particular locus (OryAAT21) being removed from the dataset. A strict analysis in which I removed OryAAT21 and all individuals with missing data for any locus yielded K=2, and these clusters showed rough correspondence to MMRWFS and BS5 in a plot of individual ancestry assignments. However, a large portion of some individual's ancestry was still assigned to the cluster more common at the other wetland complex (e.g. an individual sampled from MMRWFS having their ancestry assigned to

the cluster more common at BS5 instead of the cluster more common at MMRWFS). Since these clusters

do not perfectly correspond to geographic location, it is not clear what biologically meaningful

phenomenon they might represent. One possibility is recent admixture of two distinct populations,

perhaps interbreeding between *O. palustris* and *O. texensis* in Illinois. However, if this were the case, I

would not expect STRUCTURE to assign large amounts of an individual's ancestry to either cluster but

instead assign portions of each individual's ancestry to both clusters. Based on examining allele

frequencies of the two groups, it appears that STRUCTURE is attempting to correct the homozygote

excess at both sites by assigning individuals with homozygous genotypes to clusters where that allele is

more common, particular for OryAAT03. That suggests the assignment is in part due to excess

homozygosity in the original dataset and, ultimately, to null alleles. Determining the actual cause of the

apparent genetic structure would require further data gathering, preferably with new molecular

markers.

Rice rats in Southern Illinois, previously classified as a state-threatened species, exist in isolated

wetland habitats but share enough migrants to establish adequate levels of gene flow between them.

My sample size is relatively small, particularly in terms of the number of sites, but my results provide

some evidence that these rats are not in danger of extirpation at the northern edge of their range.

Nevertheless, there are many landscape features, such as roads, ditches, and agricultural fields, that

could affect connectivity among rice rate populations. Future research should focus on establishing

what facilitates gene flow between populations and how these features can be managed to ensure

genetic connectivity is maintained.

CHAPTER 5:

CONCLUSIONS AND POTENTIAL FUTURE RESEARCH

Rice rats appear to have been present in southern Illinois since at latest 1000 A. D. (Richards 1979)

and likely arrived earlier following a warming climate in the early Holocene (~8500 years ago, Hofman *et*

al. 1990). This northward range expansion was preceded by rice rats diverging into two distinct clades

during the Pleistocene (Hanson *et al.* 2010). Recognition of Illinois rice rats as *Oryzomys texensis* rather

than *O. palustris* adds to a developing line of evidence that the division between the two species at the

northern extent of their ranges lies further east than previously suspected. This is consistent with

Indorf's (2010) finding that samples from western Tennessee and eastern Mississippi falling out within

O. texensis. Rice rats have been recorded in central Tennessee and Kentucky (Goldman 1918; Humphrey

and Setzer 1989), but these have not been included in any phylogenetic analysis using molecular data.

Samples from these regions are crucial to locating the geographic boundary between these two species.

Study of the contact zone (if one exists) between the species would allow for a more fine-scaled

examination of their ecological relationship.

The population genetic analyses presented in chapter 4, though weakened by null alleles and

consequently inflated homozygosity, mostly support hypotheses suggested by recent studies. Cooney

(2013) showed that rice rats in southern Illinois are highly vagile with large average home ranges (3.01

ha) and 32.9 m average daily movement. He also showed that, on average, >40% of rice rat home ranges

extend outside of emergent wetland habitats. While it is important not to assume that daily movements

necessarily reflect the dispersal capabilities of a species, this might suggest an ability to disperse

between patches of habitat even in a fragmented landscape. Indeed, that dispersal ability is reflected in

the low genetic differentiation between the two wetland complexes I studied. The relatively low genetic

differentiation between MMRWFS and BS5 suggests the presence of unsampled populations that act as

stepping stones connecting the two sites. The equivalent of a single migrant per generation between

populations that are otherwise isolated results in an F_{ST} of 0.2 (Wade and McCauley 1988; Mills and Allendorf 1996). Further, the ability of rice rats to use roadside ditches (Eubanks 2009; Eubanks *et al.* 2011) might provide corridors for dispersal.

The strict STRUCTURE analysis with the OryAAT21 locus and individuals with missing data excluded recovers most individuals at MMRWFS as belonging to one genealogical group and most individuals at BS5 as belonging to another. However, the ancestry assignment of some individuals was not attributable to their source population. The population assignment also does not correspond to the mitochondrial haplotype of the sampled individuals. This apparent genetic structure may be due to two distinct populations of rats converging in Southern Illinois, but the identity of those populations cannot be discerned with my data. The pattern might be explained by interbreeding between *O. palustris* and *O. texensis*, but, for this to be true, the *O. texensis* mitochondrial haplotypes would have to have displaced any *O. palustris* haplotypes from my study sites. That would result in populations in Southern Illinois having *O. texensis* mitochondrial ancestry and mixed nuclear ancestry. Further work to determine, using nuclear markers, whether the two species are interbreeding would greatly benefit our understanding of the phylogeographic history of rice rats in the United States.

Van Der Merwe (2014) found that many rice rat populations in wetland patches at MMRWFS went extinct during a large, sustained flooding event in 2011, and that those sites were subsequently recolonized with increasing colonization rates through 2013. This may be reflected in my data, especially if populations were founded by colonists from the refuge population near the levee and then expanded. This would explain the higher level of genetic differentiation between patches at MMRWFS relative to BS5 over shorter geographic distances.

Future work on rice rats at these sites might focus on overwinter survival and reproduction. Southern Illinois is the northernmost extent of the range of *O. texensis* and overwinter survival is likely to be a key

limiting factor for rice rats at this latitude (Eubanks 2009). Further, rice rats reproduce year-round in the

southern parts of their range, and lab-reared rice rats from wild populations in Tennessee have also

been shown to reproduce year-round (Conaway 1954). It would be interesting to examine how survival

and recruitment during winter affect population sizes during spring and summer, and thus understand

how rice rats will respond to coming climate changes in Southern Illinois.

TABLES

Table 1: Primer sequences for cytochrome-b and the control region

Region	Primer	Sequence (5' → 3')	Sources
Cytochrome-b	MVZ05	CGAAGCTTGATATGAAAAACCATCGTTG	Smith and Patton 1993
	CB40RC	GCTTTGGGTGCTGRTAGTGG	Adapted from Hanson and Bradley 2008
Control Region	2340-5	GCATTTTCAGTGCTTTGC	Mendez-Harclerode *et al.* 2005
	Ory5'	TACCATGAYCTTGTAAGTC	Indorf and Gaines 2013

Table 2 Site and number of samples per hypothetical population. Hypothetical populations M2, M3, and M5 are nested within MMRWFS. Hypothetical populations B1, B2, and B3 are nested within BS5.

Site / Population		Number of samples
MMRWFS		**79**
	M2	20
	M3	32
	M5	27
BS5		**69**
	B1	20
	B2	29
	B3	20

Table 3 Microsatellite loci used in this study including primers, annealing temperature used in PCR, and the fluorescent labels used in fragment analysis.

Locus	Primers 5' → 3'	Annealing Temp. (°C)	Fluorescent Label
OryAAT03	TGGCTTCAGTGGGTATTTATTAT TGCGCACATGTATATTAAAGAA	54	HEX
OryAAT10	TTGGGTTGGCTCTAATAGAA TTTGCTATTGTTCACTTTCTAA	52	6-FAM
OryAAT21	GCCTCCTACTGTTGTTTTGTTA GCATTATCTGTTCTCTATCCACTAC	54	6-FAM
OryAAT28	TCGGGAATATAGAAAGAAGTA TGCCACTCAAAAATGATTTAAC	52	HEX
OryAAT40	GGGTGTTCCAGAGAATGAATCTA TCCCAACATCCTCACAAATTTAT	54	NED
OryAAT26	CAATTGCTATTTTTCTGTAAA CCAGCTGTTTTTTATTGTACT	50	6-FAM
OryAAT60	AAGGCAGCTAAAAAAATCTT ATCGTCTCTGTTTTTATATCAG	50	NED

Table 4: Number of alleles per locus for each of the 6 putative populations. M3, M2, and M5 are located at MMRWFS. B1, B3, and B2 are located at BS5.

Locus	M3	M2	M5	B1	B3	B2	Mean	s.d.	Total
OryAAT03	8	11	9	11	8	11	9.667	1.506	15
OryAAT10	15	11	16	10	11	14	12.833	2.483	26
OryAAT21	11	8	10	11	11	11	10.333	1.211	15
OryAAT28	8	9	9	10	8	11	9.167	1.169	14
OryAAT40	11	16	17	11	11	15	13.5	2.811	22
OryATT26	8	7	9	8	7	8	7.833	0.753	10
OryATT60	9	13	11	13	12	12	11.667	1.506	15

Table 5: Observed and expected heterozygosity (H_O / H_E) for each locus within each putative population. Significant deviations from Hardy-Weinberg Equilibrium ($p < 0.05$) are indicated with an asterisk (*).

	M3	M2	M5	B1	B3	B2
OryAAT03	0.59375 / 0.83681*	0.52632 / 0.84495*	0.36000 / 0.84735*	0.60000 / 0.86795*	0.42105 / 0.81935*	0.64286 / 0.87403*
OryAAT10	0.82143 / 0.90844*	0.64706 / 0.88057*	0.91304 / 0.93140*	0.90000 / 0.91579	0.25000 / 0.81250*	0.57143 / 0.88312*
OryAAT21	0.906 / 0.850	0.684 / 0.839*	0.593 / 0.843*	0.579 / 0.831*	0.750 / 0.867*	0.862 / 0.877*
OryAAT28	0.700 / 0.836	0.687 / 0.817	0.579 / 0.873*	0.692 / 0.886*	0.786 / 0.878	0.750 / 0.883*
OryAAT40	0.710 / 0.868	0.842 / 0.939	0.815 / 0.932*	0.611 / 0.835*	0.737 / 0.852*	0.929 / 0.910
OryATT26	0.250 / 0.832*	0.400 / 0.800*	0.467 / 0.892*	0.714 / 0.823	0.438 / 0.823*	0.500 / 0.818*
OryATT60	0.969 / 0.826	0.700 / 0.931*	0.870 / 0.904	0.944 / 0.925	0.895 / 0.899	0.862 / 0.909*

Table 6: Estimates of null allele frequency for each locus within each population using the Brookfield method (Brookfield 1996) in Micro-Checker are shown. Micro-Checker did not detect null alleles for loci marked with a dash (-).

	M3	M2	M5	B1	B3	B2
OryAAT03	0.126	0.152	0.263	0.144	0.192	0.125
OryAAT10	-	0.127	-	-	0.314	0.164
OryAAT21	-	-	0.136	0.114	-	-
OryAAT28	-	-	0.133	0.115	-	-
OryAAT40	0.078	-	-	0.104	-	-
OryATT26	0.311	0.211	0.212	-	0.186	0.173
OryATT60	-	0.127	-	-	-	-

Table 7: Pairwise Fst values for each pair of putative populations based on 7 microsatellite loci. Significant Fst values ($p < 0.05$) are marked with an asterisk (*).

Population	M3	M2	M5	B1	B3	B2
M3	0					
M2	0.025*	0				
M5	0.012*	0.000	0			
B1	0.002	0.008	0.008	0		
B3	0.031*	0.027*	0.028*	-0.004	0	
B2	0.026*	0.025*	0.011*	-0.016	-0.007	0

Table 8: Absolute number of migrants (M) between each pair of populations as estimated from F_{ST} given mutation-drift equilibrium (Slatkin 1991). M = 2nm where n is the effective population size and m is the migration rate. Populations at BS5 (B1, B2, and B2) effectively have infinite migration rates between them due to lack of meaningful genetic differentiation between them.

	M3	M2	M5	B1	B3
M2	19.29485				
M5	40.49037	1283.105			
B1	269.423	58.95807	58.97173		
B3	15.5969	18.16252	17.684	inf	
B2	18.37973	19.82968	44.48462	inf	inf

Table 9: Analysis of molecular variance (AMOVA) results. In this analysis, MMRWFS and BS5 are treated as groups with populations M3, M2, and M5 nested within MMRWFS and populations B1, B3, and B2 nested within BS5. Most of the genetic variation is explained by variation within individuals indicating very weak genetic structure. Significant fixation indices are marked with an asterisk (*).

Source of Variation	df	Variance components	Percent of variation	Fixation Index	P-value
Among wetlands	1	0.042	1.71	$F_{CT} = 0.017$	0.112 ± 0.009
Among populations within wetlands	4	0.007	0.27	$F_{SC} = 0.003$	0.463 ± 0.015
Within populations	290	2.437	98.02	$F_{ST} = 0.020$	0.000 ± 0.000*
Total	295	2.486			

Table 10: Evanno table generated by STRUCTURE Harvester (Earl and VonHoldt 2012). Mean LnP(K) is the mean log-likelihood across repetitions at a given K clusters. The Evanno method for selecting the number of real clusters across repeated STRUCTURE runs at different K uses ΔK, the second order rate of change of the likelihood function with respect to K. For the microsatellite dataset used in this study, K=2 was determined to be the best choice for number of clusters present in the dataset (underlined).

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	ΔK
1	5	-4401	0.58052	NA	NA	NA
2	5	-4069.6	1.24218	331.34	234.48	188.766
3	5	-3972.8	3.79368	96.86	27.5	7.2489
4	5	-3903.4	4.13243	69.36	44.42	10.7491
5	5	-3878.5	4.91599	24.94	34.28	6.97316
6	5	-3887.8	27.2802	-9.34	55.32	2.02784
7	5	-3952.5	37.2725	-64.66	33.18	0.8902
8	5	-3984	74.4354	-31.48	NA	NA

Table 11: Table showing ancestry assignment under K = 2 in the "red" group for individuals that were also included in the phylogeny in Chapter 3. Site represents where the geographic location the sample was taken from. "Orange" and "Blue" represent the portion of an individual's ancestry assigned to the "Orange" and "Blue" clusters in figure 7, A. "RAxML Relative" is the most closely related GenBank sequence in the RAxML majority rule consensus tree or one of the closest relatives when the closest relative is ambiguous (Figure 4). Sample names highlighted in orange have a higher portion of their ancestry assigned to the "orange" cluster. Sample names highlighted in blue have a higher portion of their ancestry assigned to the "blue" cluster.

Sample	Site	Orange	Blue	RAxML Relative
3A	MMRWFS	0.759	0.241	O. texensis TTU79152 Tennessee Shelby County 2
38	MMRWFS	0.255	0.745	O. texensis TTU79152 Tennessee Shelby County 2
37	MMRWFS	0.935	0.065	O. texensis TTU79152 Tennessee Shelby County 2
3Z	MMRWFS	0.662	0.338	O. texensis TTU79152 Tennessee Shelby County 2
3U	MMRWFS	0.07	0.93	O. texensis TTU79152 Tennessee Shelby County 2
3H	BS5	0.172	0.828	O. texensis TTU82963 Arkansas Crittenden County
3G	BS5	0.074	0.926	O. texensis TTU82963 Arkansas Crittenden County
3F	BS5	0.068	0.932	O. texensis TTU82963 Arkansas Crittenden County
3C	BS5	0.921	0.079	O. texensis TTU79152 Tennessee Shelby County 2
3B	BS5	0.877	0.123	O. texensis TTU79152 Tennessee Shelby County 2

Table 12: Ancestry assignment under K = 2 in the 'strict' analysis for individuals that were also included in the phylogeny in Chapter 3. Site represents where the geographic location the sample was taken from. "Orange" and "Blue" represent the portion of an individual's ancestry assigned to the "Orange" and "Blue" clusters in figure 7, A. "RAxML Relative" is the most closely related GenBank sequence in the RAxML majority rule consensus tree or one of the closest relatives when the closest relative is ambiguous (Figure 3). Sample names highlighted in orange have a higher portion of their ancestry assigned to the "orange" cluster. Sample names highlighted in blue have a higher portion of their ancestry assigned to the "blue" cluster.

Sample	Site	Orange	Blue	RAxML Relative
3A	MMRWFS	0.615	0.385	O. texensis TTU79152 Tennessee Shelby County 2
38	MMRWFS	0.329	0.671	O. texensis TTU79152 Tennessee Shelby County 2
37	MMRWFS	0.914	0.086	O. texensis TTU79152 Tennessee Shelby County 2
3Z	MMRWFS	0.806	0.194	O. texensis TTU79152 Tennessee Shelby County 2
3U	MMRWFS	0.097	0.903	O. texensis TTU79152 Tennessee Shelby County 2
3H	BS5	0.290	0.710	O. texensis TTU82963 Arkansas Crittenden County
3G	BS5	0.170	0.830	O. texensis TTU82963 Arkansas Crittenden County
3F	BS5	0.101	0.899	O. texensis TTU82963 Arkansas Crittenden County
3C	BS5	0.914	0.086	O. texensis TTU79152 Tennessee Shelby County 2
3B	BS5	0.868	0.132	O. texensis TTU79152 Tennessee Shelby County 2

Figure 1: Map showing approximate locations of study sites (black rectangles) from which rice rats were sampled relative to counties in Southern Illinois. Rice rats from Burning Star 5 (BS5) and the Middle Mississippi River Wetland Field Station (MMRWFS) were used in the population genetics study (Chapter 4). Rice rats from BS5, MMRWFS, and the Harrisburg, Illinois site (HAR) were used in the phylogeography study (Chapter 3).
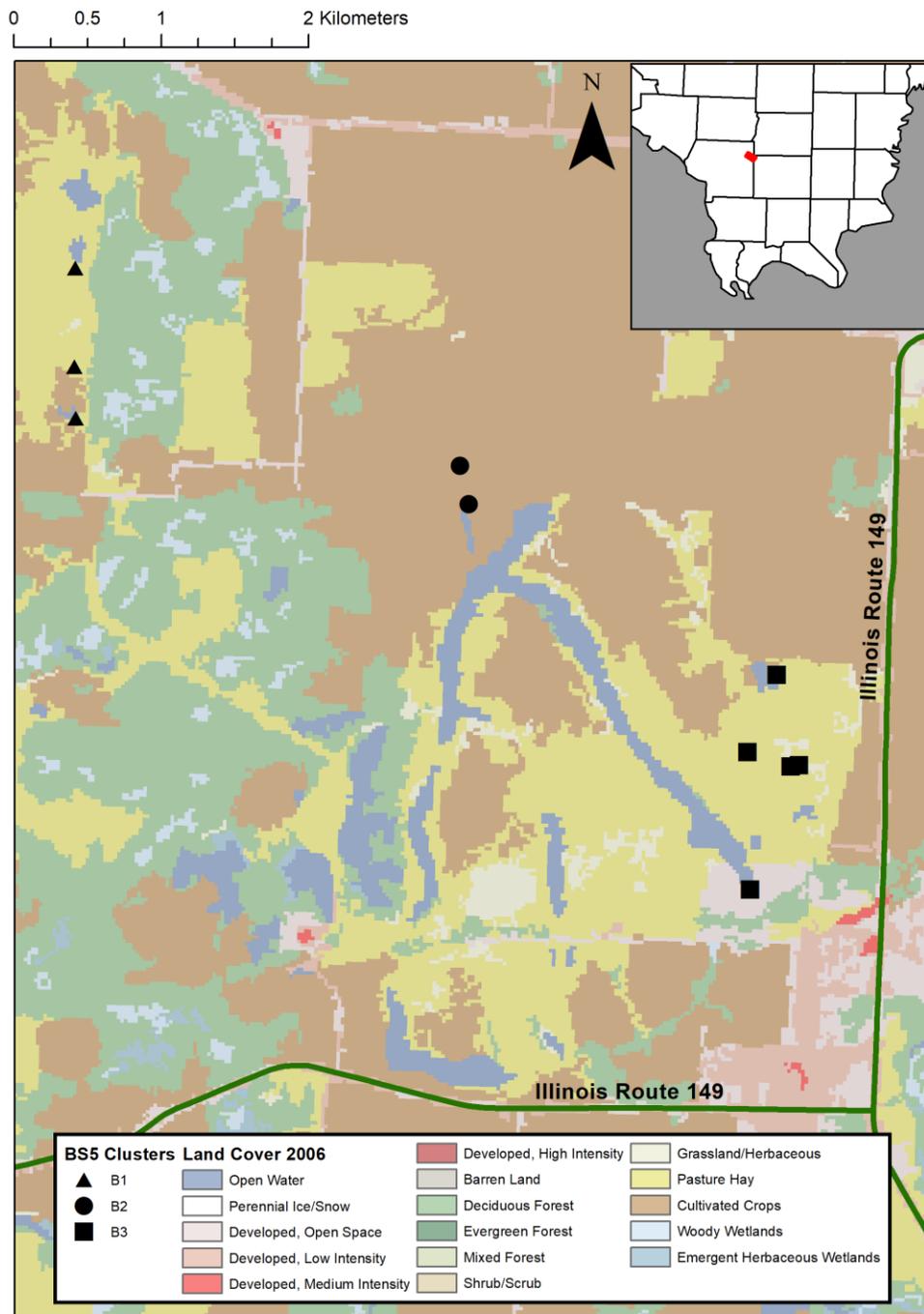
Figure 2: Land cover map of BS5 showing the locations of each trap line included in each putative population. Landcover data based on the National Land Cover Database 2006. Major roadways are shown in dark green. The inset in the upper right shows the location of this field site (red dot) relative to a map of Southern Illinois counties.
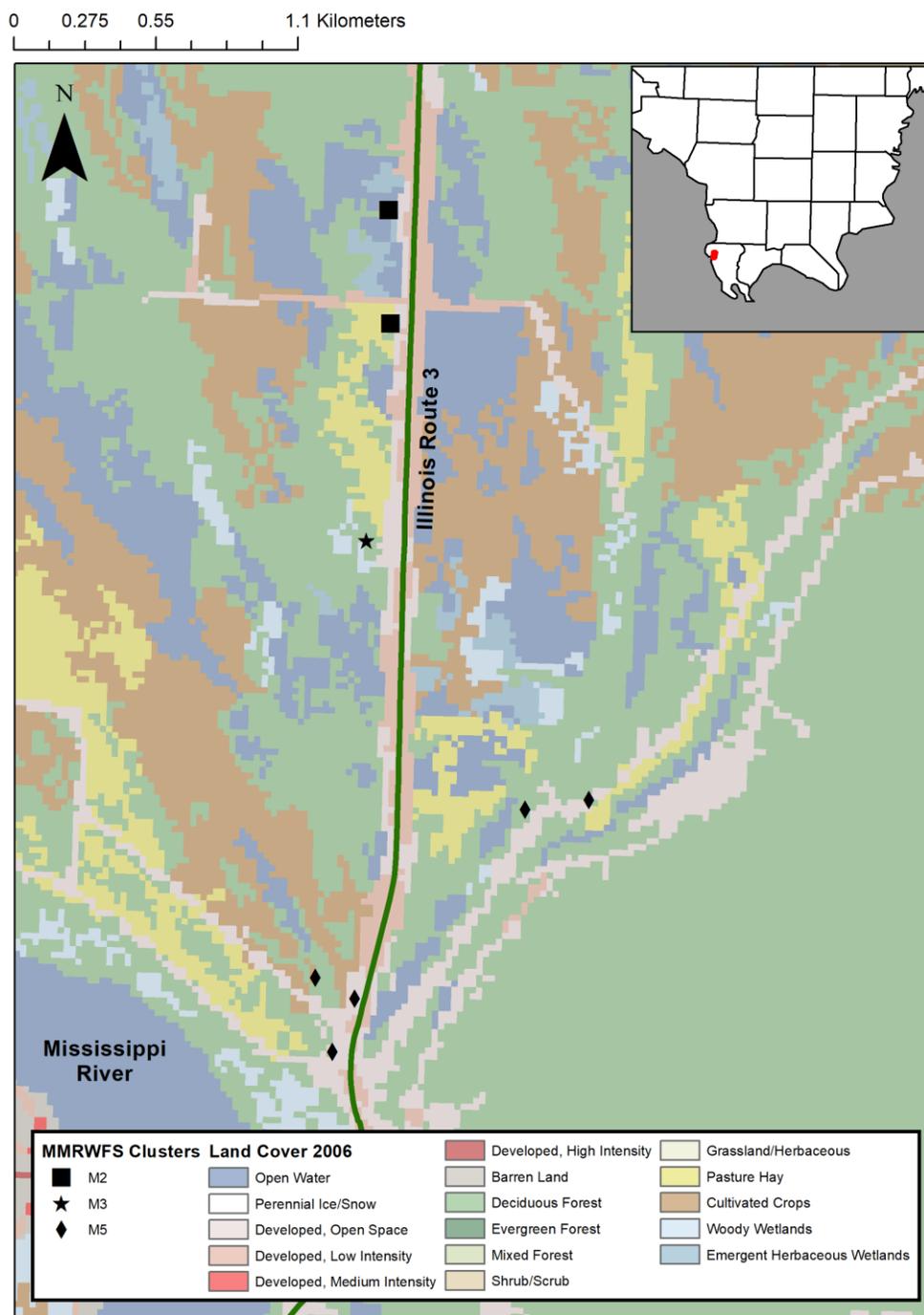
Figure 3: Land cover map of MMRWFS showing the locations of each trap line included in each putative population. Landcover data based on the National Land Cover Database 2006. Major roadways are shown in dark green. The inset in the upper right shows the location of this field site (red dot) relative to a map of Southern Illinois counties.
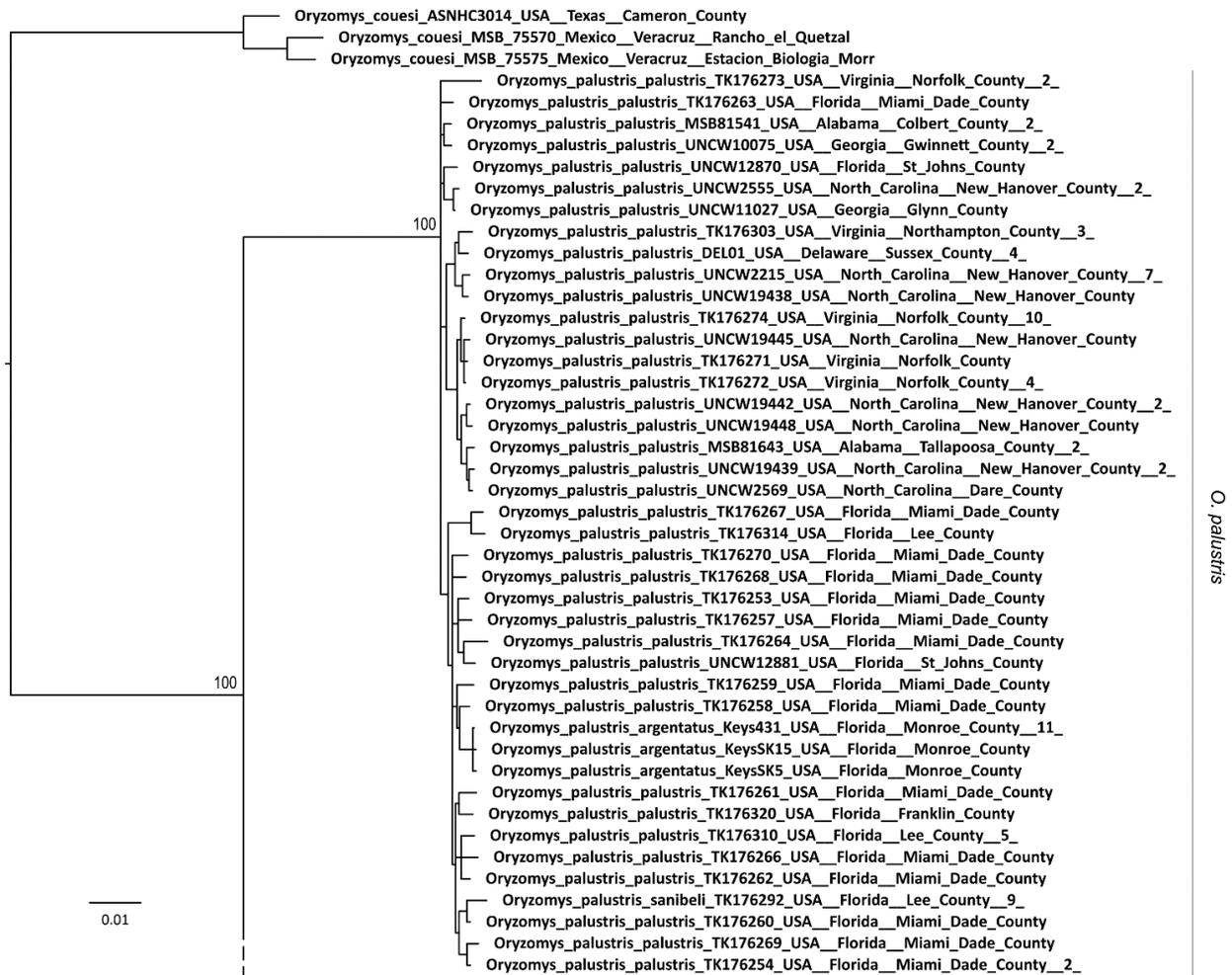
Figure 4.1: First half of the majority rule consensus tree from MrBayes analysis of concatenated cytochrome-b and control region dataset. Sequences from Illinois are highlighted in red. Support values below nodes represent Bayesian posterior probabilities (in %). Branches here are rendered relative to lengths expressed in substitutions per site. For samples from GenBank (in black), each sample ID shows the species, voucher number, and sample location in that order. The numbers after sample IDs indicate the number of duplicate haplotypes in the original GenBank PopSets. The dotted branch at the bottom of figure 4.2.

Oryzomys_texensis_TTU79152_USA__Tennessee__Shelby_County__2_
Oryzomys_texensis_TTU62979_USA__Oklahoma__Okmulgee_County
3A_USA__Illinois__Alexander_County
3B_USA__Illinois__Jackson_County
3C_USA__Illinois__Jackson_County
3U_USA__Illinois__Alexander_County
3Z_USA__Illinois__Alexander_County
38_USA__Illinois__Alexander_County
HAR1_USA__Illinois__Saline_County
HAR2_USA__Illinois__Saline_County
HAR3_USA__Illinois__Saline_County
HAR5_USA__Illinois__Saline_County
HAR6_USA__Illinois__Saline_County
37_USA__Illinois__Alexander_County
Oryzomys_texensis_LSU_M8432_USA__Louisana__Cameron_Parish__2_
Oryzomys_texensis_LSU_M8437_USA__Louisana__Cameron_Parish__2_
Oryzomys_texensis_TTU75308_USA__Texas__Freestone_County__4_
Oryzomys_texensis_TTU75307_USA__Texas__Freestone_County
Oryzomys_texensis_TTU71568_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU75193_USA__Texas__Calhoun_County
Oryzomys_texensis_TTU82870_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU71583_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71561_USA__Texas__Brazoria_County__2_
Oryzomys_texensis_TTU71592_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8460_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU78331_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8454_USA__Louisana__Cameron_Parish
Oryzomys_texensis_LSU_M8450_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU82880_USA__Texas__Galveston_County__3_
Oryzomys_texensis_TTU71559_USA__Texas__Brazoria_County__2_
Oryzomys_texensis_TTU71563_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU78332_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71595_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU82872_USA__Texas__Galveston_County__3_
Oryzomys_texensis_TTU82871_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU82881_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU82892_USA__Texas__Galveston_County
Oryzomys_texensis_TTU82882_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU71564_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8449_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU82878_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU75312_USA__Texas__Freestone_County
Oryzomys_texensis_ASNHC3031_USA__Texas__San_Patricio_County__11_
Oryzomys_texensis_UNCW3068_USA__Texas__San_Patricio_County
Oryzomys_texensis_TTU71593_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71591_USA__Texas__Brazoria_County__2_
Oryzomys_texensis_ASNHC2886_USA__Texas__Cameron_County__36_
Oryzomys_texensis_ASNHC2908_USA__Texas__Cameron_County
Oryzomys_texensis_ASNHC3001_USA__Texas__Cameron_County__8_
Oryzomys_texensis_LSU_M8453_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU71567_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8451_USA__Louisana__Cameron_Parish
Oryzomys_texensis_ASNHC2917_USA__Texas__Cameron_County__3_
Oryzomys_texensis_TTU82869_USA__Florida__Okeechobee_County__2_
Oryzomys_texensis_TTU79154_USA__Tennessee__Shelby_County
Oryzomys_texensis_TTU75163_USA__Texas__Calhoun_County
Oryzomys_texensis_ASNHC3032_USA__Texas__San_Patricio_County__2_
Oryzomys_texensis_TTU75191_USA__Texas__Calhoun_County
Oryzomys_texensis_TTU71560_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71558_USA__Texas__Brazoria_County__4_
Oryzomys_texensis_TTU62978_USA__Oklahoma__Okmulgee_County__2_
Oryzomys_texensis_LSU_M8448_USA__Louisana__Cameron_Parish
Oryzomys_texensis_LSU_M8443_USA__Louisana__Cameron_Parish__2_
Oryzomys_texensis_TTU75268_USA__Texas__Freestone_County__2_
Oryzomys_texensis_TTU82963_USA__Arkansas__Crittenden_County
Oryzomys_texensis_LSU_M8438_USA__Louisana__Cameron_Parish__3_
Oryzomys_texensis_LSU_M8455_USA__Louisana__Cameron_Parish
3D_USA__Illinois__Jackson_County
3F_USA__Illinois__Jackson_County
3G_USA__Illinois__Jackson_County
3H_USA__Illinois__Jackson_County

100

O. texensis

0.01

Figure 4.2: Bottom half of the majority rule consensus tree from MrBayes analysis of concatenated cytochrome-b and control region dataset. Continued from figure 4.1. The dotted branch at the top of the figure corresponds connects to the dotted band at the bottom of figure 4.1.

Figure 5.1: Top Half of the majority rule consensus tree from RAxML rapid bootstrapping analysis of combined cytochrome-b and control region dataset. Sequences from Illinois are highlighted in red. Support values below nodes represent the percentage of bootstrap trees recovering a clade. For samples from GenBank (in black), each sample ID shows the species, voucher number, and sample location in that order. The numbers after sample IDs indicate the number of duplicate haplotypes in the original GenBank PopSets. The branch extending off the bottom of the figure connects to the dotted branch at the top of figure 5.2.

100
64
62
90
64
92
59
97
86
62
97
77
70

Oryzomys_texensis_TTU75193_USA__Texas__Calhoun_County
Oryzomys_texensis_TTU82870_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU71568_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71583_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71592_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71561_USA__Texas__Brazoria_County__2_
Oryzomys_texensis_LSU_M8432_USA__Louisana__Cameron_Parish__2_
Oryzomys_texensis_LSU_M8437_USA__Louisana__Cameron_Parish__2_
Oryzomys_texensis_TTU75307_USA__Texas__Freestone_County
Oryzomys_texensis_TTU75308_USA__Texas__Freestone_County__4_
Oryzomys_texensis_TTU62979_USA__Oklahoma__Okmulgee_County
Oryzomys_texensis_TTU79152_USA__Tennessee__Shelby_County__2_
HAR5_USA__Illinois__Saline_County
HAR3_USA__Illinois__Saline_County
3U_USA__Illinois__Alexander_County
HAR6_USA__Illinois__Saline_County
3Z_USA__Illinois__Alexander_County
HAR1_USA__Illinois__Saline_County
37_USA__Illinois__Alexander_County
3C_USA__Illinois__Jackson_County
3A_USA__Illinois__Alexander_County
38_USA__Illinois__Alexander_County
HAR2_USA__Illinois__Saline_County
3B_USA__Illinois__Jackson_County

99
57
90
91
97
68
62
64
56
56
72
58
68
100
73
86
82
75
92
54
61
64
96
96
52

Oryzomys_texensis_LSU_M8460_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU78331_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8450_USA__Louisana__Cameron_Parish
Oryzomys_texensis_LSU_M8454_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU71563_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU82880_USA__Texas__Galveston_County__3_
Oryzomys_texensis_TTU78332_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU71559_USA__Texas__Brazoria_County__2_
Oryzomys_texensis_TTU71595_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8449_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU71564_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU82872_USA__Texas__Galveston_County__3_
Oryzomys_texensis_TTU82871_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU82882_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU82881_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU82892_USA__Texas__Galveston_County
Oryzomys_texensis_TTU82878_USA__Texas__Galveston_County__2_
Oryzomys_texensis_TTU75312_USA__Texas__Freestone_County
Oryzomys_texensis_UNCW3068_USA__Texas__San_Patricio_County
Oryzomys_texensis_ASNHC3031_USA__Texas__San_Patricio_County__11_
Oryzomys_texensis_TTU71591_USA__Texas__Brazoria_County__2_
Oryzomys_texensis_TTU71593_USA__Texas__Brazoria_County
Oryzomys_texensis_ASNHC3001_USA__Texas__Cameron_County__8_
Oryzomys_texensis_ASNHC2908_USA__Texas__Cameron_County
Oryzomys_texensis_ASNHC2886_USA__Texas__Cameron_County__36_
Oryzomys_texensis_LSU_M8453_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU71567_USA__Texas__Brazoria_County
Oryzomys_texensis_LSU_M8451_USA__Louisana__Cameron_Parish
Oryzomys_texensis_TTU79154_USA__Tennessee__Shelby_County
Oryzomys_texensis_ASNHC3032_USA__Texas__San_Patricio_County__2_
Oryzomys_texensis_TTU75191_USA__Texas__Calhoun_County
Oryzomys_texensis_TTU71560_USA__Texas__Brazoria_County
Oryzomys_texensis_TTU82869_USA__Florida__Okeechobee_County__2_
Oryzomys_texensis_TTU75163_USA__Texas__Calhoun_County
Oryzomys_texensis_ASNHC2917_USA__Texas__Cameron_County__3_
Oryzomys_texensis_TTU71558_USA__Texas__Brazoria_County__4_
Oryzomys_texensis_TTU62978_USA__Oklahoma__Okmulgee_County__2_
Oryzomys_texensis_LSU_M8448_USA__Louisana__Cameron_Parish
Oryzomys_texensis_LSU_M8443_USA__Louisana__Cameron_Parish__2_
Oryzomys_texensis_TTU75268_USA__Texas__Freestone_County__2_
Oryzomys_texensis_TTU82963_USA__Arkansas__Crittenden_County
Oryzomys_texensis_LSU_M8438_USA__Louisana__Cameron_Parish__3_
Oryzomys_texensis_LSU_M8455_USA__Louisana__Cameron_Parish
3D_USA__Illinois__Jackson_County
3F_USA__Illinois__Jackson_County
3G_USA__Illinois__Jackson_County
3H_USA__Illinois__Jackson_County

O. texensis

Figure 5.2: Bottom half of the majority rule consensus tree from RAxML rapid bootstrapping analysis of combined cytochrome-b and control region dataset. Continued from Figure 4.1. The dotted branch at the top of the figure connects to the bottom of figure 5.1.

Figure 6: STRUCTURE analysis for K=2 clusters under the admixture ancestry model. The first 79 individuals were sampled from MMRWFS while the remaining 69 individuals came from BS5. Each individual is represented by one column divided into colors representing the portion of that individual's ancestry assigned to each cluster. The colors represent STRUCTURE's ancestry assignment and do not reflect geography. Visualization was created using DISTRUCT (Rosenberg 2004)
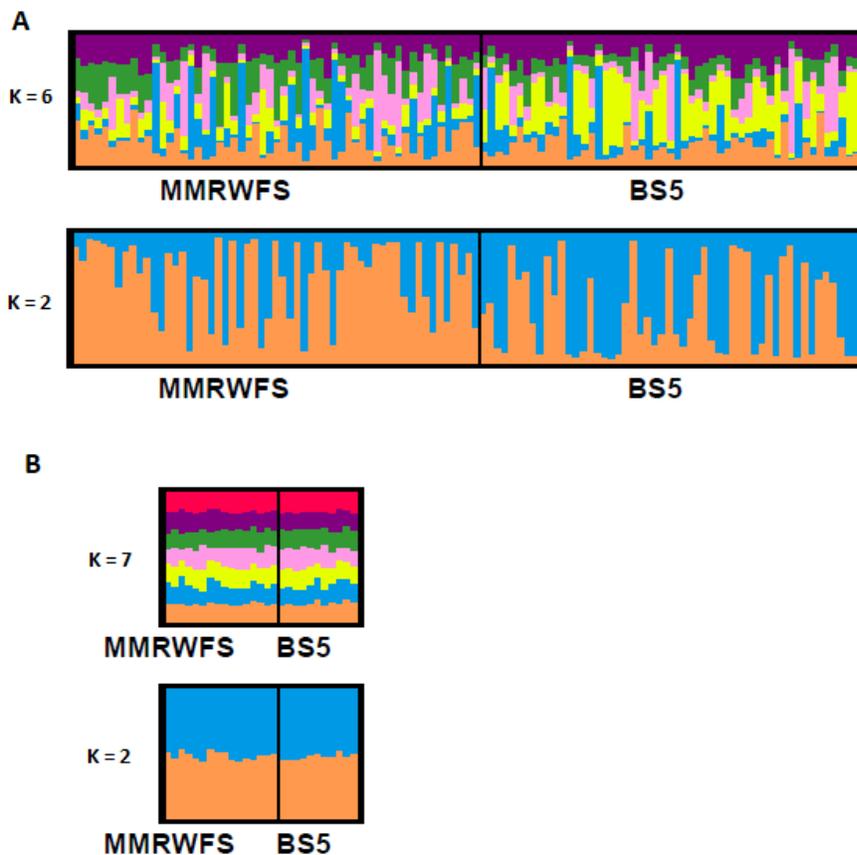
Figure 7: Separate STRUCTURE analyses for individuals with > 90% of their ancestry assigned to the "Red" cluster (A, n=110) and the "Green" cluster (B, n=27) in Figure 6. Under the Evanno method, the best number of clusters for the "Red" group was K = 6 and the best number of clusters for the "Green" group was K = 7. Visualizations of ancestry assignment at K = 2 are also included for comparison to the two sampling sites. Each individual is represented by one column divided

No Loci Removed
K = 2

OryAAT21 Removed
K = 5

OryAAT21 Removed
K = 2

OryAAT28 Removed
K = 2

OryAAT60 Removed
K = 2

OryAAT26 Removed
K = 2

OryAAT10 Removed
K = 2

OryAAT40 Removed
K = 2

OryAAT03 Removed
K = 2
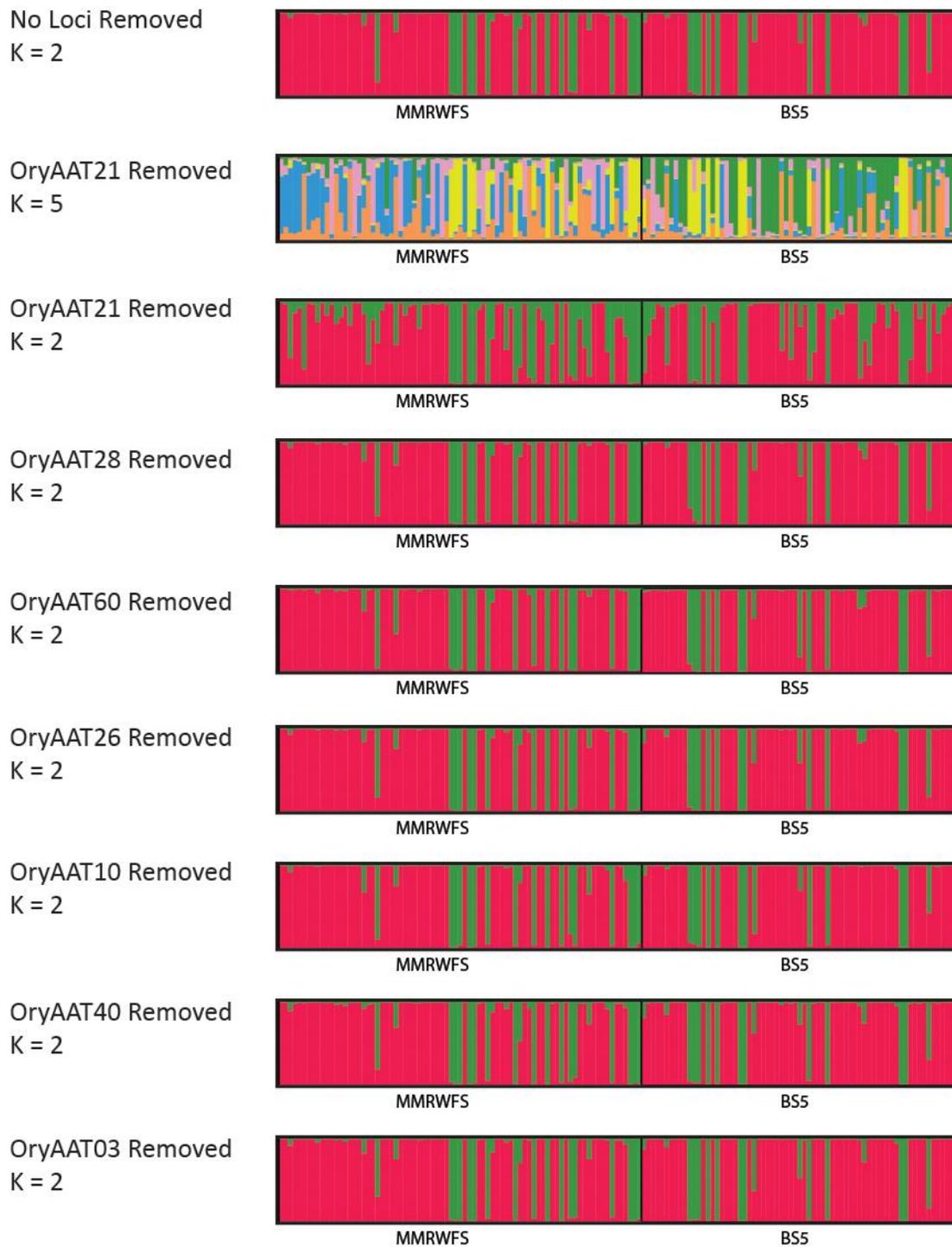
Figure 8: Separate STRUCTURE analyses for reduced data sets in which one microsatellite locus was removed. Each analysis is visualized assuming the optimal number of clusters under the Evanno method. K = 2 for the unreduced dataset is included at the top for comparison. For the OryAAT21 removed dataset, K = 2 is also included for comparison. For each analysis, each individual is represented by one column divided into colors representing the portion of that individual's ancestry assigned to each cluster.
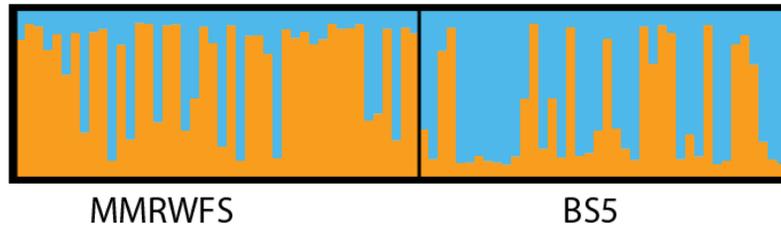
Figure 9: STRUCTURE analysis for K =2 clusters using a reduced dataset with no missing data and the OryAAT21 locus removed. Each individual is represented by one column divided into colors representing the portion of that individual's ancestry assigned to each cluster.

REFERENCES

Abuzeineh, A.A., R.D. Owen, N.E. Mcintyre, C.W. Dick, E. Richard, T. Holsomback, S. Url, D. Owen, and W. Dick. 2007. Response of marsh rice rat (Oryzomys palustris) to inundation of habitat. The Southwestern Naturalist, 52, 75–78.

Allendorf, F.W., G.H. Luikart, and S.N. Aitken. 2012. Conservation and the Genetics of Populations, 2nd ed. Wiley-Blackwell, Hoboken, NJ.

Altekar, G., S. Dwarkadas, J.P. Huelsenbeck, and F. Ronquist. 2004. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics, 20, 407–415.

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. 1998. The history and purview of phylogeography. Molecular Ecology, 371–379.

Avise, J.C. 2001. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Bradley, R.D. and R.J. Baker. 2001. A Test of the Genetic Species Concept: Cytochrome-b Sequences and Mammals. Journal of Mammalogy, 82, 960–973.

Brookfield, J.F. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol, 5, 453–455.

Church, S.H., J.F. Ryan, and C.W. Dunn. 2015. Automation and evaluation of the SOWH test with SOWHAT. Systematic Biology, 64, 1048–1058.

Conaway, C.H. 1954. The Reproductive Cycle of Rice Rats (Oryzomys palustris palustris) in Captivity. Journal of Mammalogy, 35, 263–266.

Cooney, S.A. 2013. Landscape Permeability and Home Range Composition of the Marsh Rich Rat (Oryzomys Palustris) in Southern Illinois. Southern Illinois University.

Cornuet, J.M. and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics, 144, 2001–14.

Cory, C.B. 1912. The Mammals of Illinois and Wisconsin. Field Museum Natural History Zoological Series, 11, 1–505.

Earl, D.A. and B.M. VonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, 4, 359–361.

Edger R, C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32, 1792–1797.

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

Estoup, A., P. Jarne, and J.-M. Cornuet. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Molecular Ecology, 11, 1591–1604.

Eubanks, B.W. 2009. Status and Distribution of the State-Threatened Marsh Rice Rat (Oryzomys Palustris) in Illinois. Southern Illinois University.

Eubanks, B.W., E.C. Hellgren, J.R. Nawrot, and R.D. Bluett. 2011. Habitat associations of the marsh rice rat (Oryzomys palustris) in freshwater wetlands of southern Illinois. Journal of Mammalogy, 92, 552–560.

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, 2611–20.

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, 491, 479–491.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary bioinformatics online, 1, 47–50.

Forys, E. and R. Dueser. 1993. Inter-island movements of rice rats (Oryzomys palustris). American Midland Naturalist, 130, 408–412.

Gelman, A. and D.B. Rubin. 1992. Inference from Iterative Simulation Using Multiple Sequences. Statistical Science, 7, 457–511.

Goldman, E.A. 1918. The Rice Rats of North America:(Genus Oryzomys).

Guindon, S., J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology, 59, 307–321.

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.

Hanson, D.J. and R.D. Bradley. 2008. Molecular diversity within Melanomys caliginosus (rodenti: Oryzomyini): Evidence for multiple species. Occacional papers, 2–11.

Harlan, R. 1837. Description of a new species of Quadruped, of the order Rodentia, inhabiting the United States. American Journal of Science and Arts, 31, 385–386.

Harrison, R.G. and E.L. Larson. 2014. Hybridization, introgression, and the nature of species boundaries. Journal of Heredity, 105, 795–809.

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.

Hofman, J., J. Gardner, and M. Morris. 1990. Distribution, abundance, and habitat of the marsh rice rat (oryzomys palustris) in southern Illinois. Transactions of the Illinois State Academy of Science, 83, 162–180.

Hofmann, J.E. and J.E. Gardner. 1992. A Radiotelemetric Study of Home Range Size and Dispersal of the Marsh Rice Rat (Oryzomys Palustris) in Southern Illinois: Progress Report.

Holsinger, K.E. and B.S. Weir. 2009. Genetics in geographically structured populations: defining, estimating and interpreting F(ST). Nature reviews. Genetics, 10, 639–50.

Howell, A.H. 1910. Notes on the mammals of the Middle Mississippi Valley, with description of a new woodrat. Proceedings of the Biological Society of Washington, 23, 23–34.

Hubisz, M.J., D. Falush, M. Stephens, and J.K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular ecology resources, 9, 1322–32.

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

Humphrey, S.R. and H.W. Setzer. 1989. Geographic Variation and Taxonomic Revision of Rice Rats (Oryzomys palustris and O. argentatus) of the United States. Journal of Mammalogy, 70, 557–570.

Indorf, J.L. 2010. Phylogeography of the Marsh Rice Rat (Oryzomys Palustris) in Wetlands of the Southeastern United States. University of Miami.

Indorf, J.L. and M.S. Gaines. 2013. Genetic divergence of insular marsh rice rats in subtropical Florida. Journal of Mammalogy, 94, 897–910.

Jaarola, M., P. Vogel, M. Beysard, G. Heckel, and N. Perrin. 2011. Asymmetric and differential gene introgression at a contact zone between two highly divergent lineages of field voles (Microtus agrestis). Journal of Evolutionary Biology, 25, 400–408.

Jakobsson, M. and N. a Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics (Oxford, England), 23, 1801–6.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28, 1647–1649.

Lanfear, R., P.B. Frandsen, A.M. Wright, T. Senfeld, and B. Calcott. 2016. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Molecular Biology and Evolution, msw260.

Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the ESA, 237–240.

Luikart, G., F.W. Allendorf, J.M. Cornuet, and W.B. Sherwin. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. The Journal of heredity, 89, 238–47.

Maestri, R., N.S. Upham, and B.D. Patterson. 2018. Tracing the diversification history of a Neogene rodent invasion into South America. Ecography, 1–13.

Mathias, M. da L., G. Heckel, C. Bastos-Silveira, S.M. Santos, and R. Monarca. 2012. Deep mitochondrial introgression and hybridization among ecologically divergent vole species. Molecular Ecology, 21, 5309–5323.

Mclaughlin, C.A. and W.B. Robertson. 1951. A New Record of the Rice Rat, Oryzomys palustris palustris, from Southern Illinois. Natural History Miscellanea, 80, 14–15.

Mendez-Harclerode, F.M., J.D. Hanson, C.F. Fulhorst, M.L. Milazzo, D.C. Ruthven, and R.D. Bradley. 2005. Genetic diversity within the southern plains woodrat (Neotoma micropus) in southern Texas. Journal of Mammalogy, 86, 180–190.

Van der Merwe, J. 2014. Ecology of the Marsh Rice Rat (Oryzomys Palustris) in Southern Illinois: Wetland Dynamics, Metapopulations, and Trophic Position. Southern Illinois University.

Van der Merwe, J., E.C. Hellgren, and E.M. Schauber. 2016. Variation in metapopulation dynamics of a wetland mammal: The effect of hydrology Peters, D.P.C., editor. Ecosphere, 7.

Miller, M.A., W. Pfeiffer, and T. Schwartz. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *in* 2010 Gateway Computing Environments Workshop (GCE), pp. 1–8. IEEE.

Mills, L.S. and F.W. Allendorf. 1996. The One-Migrant-per-Generation Rule in Conservation and Management. Conservation Biology, 10, 1509–1518.

Necker, W.L. and D.M. Hatfield. 1951. Mammals of Illinois. Bulletin of the Chicago Academy of Sciences, 9, 17–60.

Negus, N., E. Gould, and R. Chipman. 1961. Ecology of the rice rat, Oryzomys palustris (Harlan), on Breton Island, Gulf of Mexico, with a critique of the social stress theory. Tulane Studies in Zoology, 8, 95–123.

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

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.

Park, S.D.E. 2001. Trypanotolerace in West African Cattle and the Population Genetic Effects of Selection. Trinity College Dublin.

Piry, S., G. Luikart, and J.M. Cornuet. 1999. BOTTLENECK: a computer program for detecting recent reduction 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.

Rambaut, A. and N.C. Grass. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Computer Applications in the Biosciences, 13, 235–238.

Richards, R.L. 1979. Rice Rat (Oryzomys palustris) Remains from Southern Indiana Caves. Proceedings of the Indiana Academy of Sciences, 89, 425–431.

Ronquist, F. and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19, 1572–1574.

Rosenberg, N.A. 2004. DISTRUCT: A program for the graphical display of population structure. Molecular Ecology Notes, 4, 137–138.

Sikes, R.S., W.L. Gannon, and T.A.C. and U.C. of the A.S. of Mammalogists. 2007. Guidelines of the American Society of Mammalogists for the use of wild mammals in research. Journal of Mammalogy, 88, 809–823.

Slatkin, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. Theoretical Population Biology, 12, 253–262.

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

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.

Stamatakis, A., P. Hoover, and J. Rougemont. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Systematic Biology, 57, 758–771.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30, 1312–1313.

Wade, M.J. and D.E. McCauley. 1988. Extinction and recolonization : their effects on the genetic differentiation of local populations. Evolution, 42, 995–1005.

Wang, Y., C.R. Hughes, E.A. Gines-Candelaria, and M.S. Gaines. 2000. Polymorphic microsatellite loci of Oryzomys palustric, 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.

Weir, B.S. 1996. Genetic Data Analysis II: Methods for Discrete Population Genetic Data.

Weir, B.S. and C.C. Cockerham. 1984. Estimating F-Statistics for the Analysis of Population Structure. Evolution, 38, 1358.

Wolfe, J. 1982. Oryzomys palustris. Mammalian Species, 176, 1–5.

Wright, S. 1931. Evolution in Mendelian populations. Genetics, 16, 97–159.

APPENDICES

APPENDIX A

NEWLY OBTAINED RICE RAT CYTB SEQUENCES IN FASTA FORMAT.

>3A_USA:_Illinois,_Alexander_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3B_USA:_Illinois,_Jackson_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3C_USA:_Illinois,_Jackson_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3D_USA:_Illinois,_Jackson_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATTTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACACATATTTGCCGAGACGTAAACTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACGACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATCACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAGACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATATTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAACTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3F_USA:_Illinois,_Jackson_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATTTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACACATATTTGCCGAGACGTAAACTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACGACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATCACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAGACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATATTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAACTCCTATACTGA

ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA
>3G_USA:_Illinois,_Jackson_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATTTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACACATATTTGCCGAGACGTAAACTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACGACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATCACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAGACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATATTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAACTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3H_USA:_Illinois,_Jackson_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATTTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACACATATTTGCCGAGACGTAAACTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACGACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATCACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAGACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATATTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAACTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3U_USA:_Illinois,_Alexander_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>3Z_USA:_Illinois,_Alexander_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>37_USA:_Illinois,_Alexander_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATACCCTAAAACTACACGTA

>38_USA:_Illinois,_Alexander_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>HAR1_USA:_Illinois,_Saline_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>HAR2_USA:_Illinois,_Saline_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>HAR3_USA:_Illinois,_Saline_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>HAR5_USA:_Illinois,_Saline_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

>HAR6_USA:_Illinois,_Saline_County
ACACCCACTACTAAAAATTATCAACCACTCATTTATTGACCTACCAACCCCATCTAACATCTCATCATGATGAAACTTCGGATCACTACTAGGAATCTGCCTCATAG
TACAAATCATCACAGGACTATTCCTAGCAATACATTACACATCAGATACAACCACAGCATTCTCTTCAGTTACCCATATTTGCCGAGACGTAAATTACGGTTGACT
TATTCGATATGCCCACGCTAACGGAGCCTCAATATTCTTTATCTGCCTCTTCATCCACGTCGGACGAGGTATATACTATGGATCATTTATACTCAACGAAACCTGA
AACATCGGAATCATTCTACTACTTACAACTATAGCAACAGCATTCGTAGGTTATGTCCTACCATGAGGACAAATATCATTCTGAGGAGCTACAGTAATCACCAAC
CTACTCTCAGCAATCCCATACATCGGAACTACCCTAGTTGAATGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCGCATTCCACTTTA
TCCTTCCATTTATTATTACAGCCTTAGTTCTAGTACACCTTCTATTTCTTCATGAAACAGGATCAAACAACCCCTCAGGACTAAACTCAAACTCAGACAAAATTCCA
TTCCACCCATACTACACAATAAAAGACCTATTAGGAATTCTACTCCTATTAATGGTTCTCATATTTATGGTTTTATTTTTCCCAGATGTTCTCGGAGACCCAGATAA
TTATACACCAGCAAATCCACTCAACACCCCAGCACACATTAAGCCAGAATGATACTTCCTCTTCGCATACGCCATTCTACGATCTATCCCAAACAAATTAGGAGGG
GTATTAGCCCTCCTACTCTCAATTCTTATCCTAGCTACATTACCACTTCTCAACTTCTCCAAACAACAAGGATTAACCTACCGCCCAATCACTCAATTCCTATACTGA
ATTTTCGTAGCCAACCTTCTCATCTTAACATGAATTGGAGGTCAACCAGTAGAATACCCATTCACAATAATCGGACAAATCTCATCAATCCTATATTTTACTATTAT
TGTTATTCTTATACCAATTGCTAGCATAATTGAAAACAATATCCTAAAACTACACGTA

APPENDIX B

NEWLY OBTAINED RICE RAT CONTROL REGION SEQUENCES IN FASTA FORMAT.

>3A_USA:_Illinois,_Alexander_County
GAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTTTA
CATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCACAC
CATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACCTC
AACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGC
CCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTAT
CGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTG
TATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAGT
CAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>3B_USA:_Illinois,_Jackson_County
GAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTTTACA
TTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCACACCA
TAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACCTCAA
CCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGCCC
ACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTATCG
TCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTGTA
TTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAGTC
AATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAGT
GGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACAC

>3C_USA:_Illinois,_Jackson_County
GAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTTTA
CATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCACAC
CATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACCTC
AACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGC
CCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTAT
CGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTG
TATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAGT
CAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>3D_USA:_Illinois,_Jackson_County
GGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTACATAGTACATCATACATAACTATGTACATTTT
ACATTAAATTATTCCCCATTATCATATAAGCTAGCAATACAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTCAACTATACCAAACATGCATATCCCCAC
ACCATAAAATCTATGACTATAAGACATCCTTATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTACACTAACC
TCAACCATACATAAAACATATCAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCC
GCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTT
ATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGG
TGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCA
ATCAATGGTCCAGCGGACAATCATTTAATGTTAGTAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCA
AGTGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>3F_USA:_Illinois,_Jackson_County
GAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTACATAGTACATCATACATAACTATGTACATTTTA
CATTAAATTATTCCCCATTATCATATAAGCTAGCAATACAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTCAACTATACCAAACATGCATATCCCCACAC
CATAAAATCTATGACTATAAGACATCCTTATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTACACTAACCTC
AACCATACATAAAACATATCAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGC
CCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTAT
CGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTG
TATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAAT
CAATGGTCCAGCGGACAATCATTTAATGTTAGTAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTT

>3G_USA:_Illinois,_Jackson_County
GAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTACATAGTACATCATACATAACTATGTACATTTTA
CATTAAATTATTCCCCATTATCATATAAGCTAGCAATACAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTCAACTATACCAAACATGCATATCCCCACAC
CATAAAATCTATGACTATAAGACATCCTTATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTACACTAACCTC
AACCATACATAAAACATATCAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGC
CCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTAT

CGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTG
TATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAAT
CAATGGTCCAGCGGACAATCATTTAATGTTAGTAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>3H_USA:_Illinois,_Jackson_County
AGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTACATAGTACATCATACATAACTATGTACATTTTAC
ATTAAATTATTCCCCATTATCATATAAGCTAGCAATACAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTCAACTATACCAAACATGCATATCCCCACACC
ATAAAATCTATGACTATAAGACATCCTTATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTACACTAACCTCA
ACCATACATAAAACATATCAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGCC
CACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTATC
GTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTGT
ATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAATC
AATGGTCCAGCGGACAATCATTTAATGTTAGTAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAGT
GGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>3U_USA:_Illinois,_Alexander_County
GAAGGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACA
TTTTACATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCC
CACACCATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTA
ACCTCAACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAA
CCCGCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATG
ATTTATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTG
GGGTGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCT
CTCAGTCAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCA
TTCAAGTGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>3Z_USA:_Illinois,_Alexander_County
AGGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTT
TACATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCAC
ACCATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACC
TCAACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCC
GCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTT
ATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGG
TGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCA
GTCAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCA
AGTGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>37_USA:_Illinois,_Alexander_County
GAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTTTA
CATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCACAC
CATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACCTC
AACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGC
CCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTAT
CGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTG
TATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAGT
CAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATT

>38_USA:_Illinois,_Alexander_County
GAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTTTA
CATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCACAC
CATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACCTC
AACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGC
CCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTAT
CGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTG
TATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAGT
CAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>HAR1_USA:_Illinois,_Saline_County
AGGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTT
TACATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCAC
ACCATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACC
TCAACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCC
GCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTT

ATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGG
TGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCA
GTCAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCA
AGTGGTTCAAAAATTATGACTTAAATTTTAG

>HAR2_USA:_Illinois,_Saline_County
AGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACATTTTAC
ATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCCCACACC
ATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTAACCTCA
ACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAACCCGCC
CACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATGATTTATC
GTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTGGGGTGT
ATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCTCTCAGT
CAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCATTCAAG
TGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>HAR3_USA:_Illinois,_Saline_County
GAAGGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACA
TTTTACATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCC
CACACCATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTA
ACCTCAACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAA
CCCGCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATG
ATTTATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTG
GGGTGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCT
CTCAGTCAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCA
TTCAAGTGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>HAR5_USA:_Illinois,_Saline_County
GAAGGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACA
TTTTACATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCC
CACACCATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTA
ACCTCAACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAA
CCCGCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATG
ATTTATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTG
GGGTGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCT
CTCAGTCAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCA
TTCAAGTGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

>HAR6_USA:_Illinois,_Saline_County
GAAGGAGAGGAAACCCCCCACTACCAGCACCCAAAGCTGACATTCTTCCAATTAAATTACTTCTTGTACATAATTTTATATAGTACATCATACATAACTATGTACA
TTTTACATTAAATTATTCCCCATTATCATATAAGCTAGCAATATAATTCAATTATTATCCACCCAAAATGTTTAATCAACATTAAACTATATCAAACATGCATATCCC
CACACCATAAAATCTATGACTACAAGACATCTATATGTTTAATCAACATTATACACTTACCCACATAACTATTATCCTATACTAATTATTAAACCATAACCTATACTA
ACCTCAACCATACATAAAACATAACAGTCACGATCAATCCCAGTCCAAATGACTATCCCCTACTACATGTGGTCCCTTAATCTACCATCCTCCGTGAAATCAGCAA
CCCGCCCACCTAGACGGCTCTTCTTGCTCTGAGCCCATAAAATCAGGGGGGTGACTAACCTGAAACTTTATCAGGCATCTGGTTCTTACTTCAGGGCCATCAAATG
ATTTATCGTCCATACGTTCCCCTTAAATAAGACATCACGATGGTGTAGGTCTAACCTCTCGTTACCCAACATCCATTGGTTCCTTACATTTAGTAGTTTTTATTTTTG
GGGTGTATTACTCAACACAGCCGTCAAGGCATGAAGGTCAGCTTTACGTCAAGCCAGTCATCAAGTTAAGGATCATTTATCAACACAACCAAATCATCTTAAGCT
CTCAGTCAATGGTCCAGCGGACAATCATTTAATGTTAATAAGACATAATAATTTTATACCATCATCCTTTTCACCCATTACCCCTCCAATCACACTCACTCTCATCCA
TTCAAGTGGTTCAAAAATTATGACTTAAATTTTAGTATTGGGAAAAAAATCCAATATCCAACCAACCATACATCATTTTATTAATATTCTTCTACTGTTCAACACA

VITA

Graduate School

Southern Illinois University

Conrad Williams

conrad.williams@siu.edu

University of Texas at Austin

Bachelor of Science in Biology (Ecology, Evolution, and Behavior), May 2011

Thesis Title:

Population Genetics of Rice Rats (Oryzomys palustris) at the Northern Edge of the Species Range

Major Professor: Dr. Kamal Ibrahim