Fall 11-15-2016

The molecular systematics and phylogeography of the widespread North American meadow vole (Microtus pennsylvanicus)

Donavan J. Jackson
University of New Mexico - Main Campus

Donavan J. Jackson
Candidate

Biology
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Joseph A. Cook , Chairperson

Michael J. Andersen

Stephen Greiman

The molecular systematics and phylogeography of the widespread North American meadow vole (*Microtus pennsylvanicus*)

by

DONAVAN J. JACKSON

B.S., Biology, University of New Mexico, 2013

THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Biology

University of New Mexico

Albuquerque, New Mexico

December, 2016

ACKNOWLEDGEMENTS

First and foremost, I want to thank my advisor Joe Cook for all of his help, guidance and encouragement throughout my time at UNM as an undergraduate and graduate student. I have Joe to thank for many invaluable experiences and memories that have allowed me to travel and enjoy fieldwork in Canada, Panama, Alaska and Mongolia. Furthermore, it was the encouragement I received from Joe that pushed me to continue research. I would also like to thank my committee members, Michael Andersen and Stephen Greiman, for their input that greatly improved earlier versions of my thesis. I cannot thank the Cook Lab enough for the guidance and help that they provided these past few years, but more importantly I want to thank you all for your friendship and patience. I would especially like to thank Kayce Bell and Bryan McLean. Thank you both for helping me become the mammalogist that I am today. Also, thank you to my family for supporting me and giving me the encouragement I needed. Finally, I extend my gratitude to all the natural history collections and museum curators that kindly provided the specimens that made this study possible. I want to specially thank the Florida Museum of Natural History and their help with acquiring samples in Florida. Lastly, I want to extend my thanks to all the generous funding sources that supported my education and research, especially the UNM Biology Department, UNM Graduate and Professional Student Association, Indigenous STEM Research and Graduate Education program and the Gates Millennium Scholars Program.

The molecular systematics and phylogeography of the widespread North American meadow vole (*Microtus pennsylvanicus*)

by

Donavan J. Jackson

B.S., Biology, University of New Mexico 2013
M.S., Biology, University of New Mexico 2016

ABSTRACT

The climatic and environmental fluctuations of the Quaternary played an integral role in geographic distribution and genetic structure within many organisms today. Understanding how these historical biogeographic events may have partitioned genetic variation throughout the landscape is critical to forecasting the implications of modern climate change and how animals will respond to projected climate shifts. Through geographic and taxonomically comprehensive sampling, we used multi-locus and species delimitation analyses along with niche modeling methods to investigate the evolutionary and biogeographic history of the meadow vole, *Microtus pennsylvanicus*. The findings of this study highlight the vulnerability and importance that special consideration and conservational strategies be taken to preserve peripheral populations that harbor unique characteristics. The dynamic evolutionary and geographic history of the meadow vole furthers our understanding of the important role Quaternary climatic oscillations played in the diversification and geographic distribution of organisms today.

Table of Contents

Chapter 1

Thesis Introduction

The cyclical climatic and environmental events of the Pleistocene (2.6 – 0.012 mya) played a key role in shaping the contemporary distribution of plants and animals in North America, leaving genetic signatures that may be used to infer spatial and temporal impacts of this dynamic epoch (Hofreiter & Stewart 2009). Contemporary studies utilizing molecular data aim to better understand the impact of these climatic fluctuations and their role in partitioning genetic variation across the landscape (Avise *et al.* 1987; Brunsfeld *et al*. 2000; Soltis *et al*. 2006). Using molecular techniques, for example, it is possible to identify geographically structured lineages and infer glacial refugia, barriers preventing gene flow, areas that have been recently colonized, and secondary contact between divergent populations (e.g. Dragoo *et al*. 2006; Kohli *et al.* 2015). Moreover, molecular studies are refining our understanding of global biodiversity as new mammalian species are described at a fairly constant rate annually (Reeder et al. 2007). Historically, taxonomy was largely based on morphological characteristics, but those studies often failed to identify diversity, termed "cryptic" diversity, that molecular studies are now revealing. Furthermore, enhancing our knowledge of how organisms responded to historical climate change may provide vital information to ensure that effective actions are taken to preserve all biodiversity throughout North America.

Studying wide-ranging organisms provides the opportunity to investigate how a single species responded to the glacial fluctuations of the Pleistocene. Often, species have been shown to have persisted in a single refugium (Bidlack & Cook 2001; Brunhoff *et al.* 2003) whereas more widespread species sometimes persisted in multiple refugia (Aubry

et al. 2009). Additionally, post-glacial dynamics often varied between species based on physiological, geographical or environmental factors that collectively affected an organism's ability to move throughout the landscape temporally and spatially (DeChaine 2008). During the glacial-interglacial oscillations throughout this time period, organisms constantly expanded and retreated to independent refugia where they persisted in isolation, leaving distinctive genetic patterns. Using molecular data, studies are now revealing significant geographic structure (Dragoo *et al.* 2006) and identifying cryptic species (Fennessy *et al*. 2016) , thus elucidating the limits of antiquated taxonomy (Hope *et al.* 2010) and aiding in the development of new management strategies (Malaney & Cook 2013). Additionally, studies are beginning to show that low-latitude populations often disproportionately harbor unique genetic diversity that is important to the phylogenetic history and evolutionary potential of species that historically experienced poleward shifts in their distributions (Hampe & Petit 2005). Herein, we addressed the current taxonomy within the meadow vole, *Microtus pennsylvanicus*, using mitochondrial and multilocus analyses to characterize genetic variation within this species. Subsequently, we focused on peripheral populations along the southern edge of this species to emphasize some of the implications modern climate change may have for geographically restricted populations. Finally, we investigated the phylogeography of this wide-ranging species with range-wide sampling using demographic and multilocus analyses to better understand the evolutionary history of the meadow vole.

In Chapter 2, I re-evaluate the taxonomy of the meadow vole (*Microtus pennsylvanicus*) with range-wide sampling and assess the implications of modern climate change on peripheral populations. This vole has the widest range of any vole in North

America and has isolated populations along the western, southern and eastern range limits. Specimens were sampled trans-continentally from Alaska to Newfoundland and from the Arctic Ocean to coastal Florida. *Microtus pennsylvanicus* is commonly associated with mesic meadows throughout the coniferous, deciduous and boreal forest, but is also found in coastal marshes and tundra. Presumed to be the most widespread microtine during the Pleistocene (Kurten & Anderson, 1980), we used this wide-ranging organism to: 1) examine concordance between the morphologically derived subspecies and genetic data; 2) evaluate the implications of past extirpation and future projections of persistence for populations along the southern periphery; and 3) use multilocus species delimitation methods to determine if current taxonomy reflects genetic data. Using mitochondrial data, my analyses strongly supported 4 distinctive genetic lineages across this species. Multilocus species delimitation analyses support the presence of multiple incipient species that we recommend warrant elevation to species. Finally, special management strategies need to be taken to ensure the persistence of the putative coastal species, *M. dukecampbelli*, in light of the precarious future of this species due a projected rise in sea level and increased frequency of tropical storms (Hotaling *et al*. 2010).

In Chapter 3, I assess the phylogeographic history of *M. pennsylvanicus* using extensive range-wide sampling, demographic analyses and niche modeling methods. Previous studies of species within the genus *Microtus* have shown that genetic signatures of isolation and expansion are evident due to rapid genetic changes that accumulate in *Microtus* species (Jaarola *et al.* 2004). The wide latitudinal and longitudinal distributional span of this species provides the opportunity to test for the potential occupation and influence of persistence in multiple refugia, subsequent post-glacial colonization

dynamics following the retreat of ice sheets, and the projected effects of modern climate change. My analyses of this species illustrates how the dynamic biogeographic history of North America greatly influenced both genetic structure and geographic distribution in the past and currently. This study supports the hypothesis that the meadow vole occupied at least 4 ancestral areas or glacial refugia during the Last Glacial Maximum that resulted in 4 genetically distinct lineages that subsequently rapidly expanded following glacial retreat.

By studying this species, I hope to further our knowledge of the complex and integral role that the Pleistocene climatic fluctuations had on the contemporary genetic structure and distribution of North American species. This study provides deeper resolution and a comprehensive understanding of the systematics and biogeographic history of this widespread organism. Additionally, I will provide data that can be utilized in future comparative studies to improve our knowledge of the dynamics of the biogeographic and evolutionary history of organisms within North America. Furthermore, this study provides a basis for the effective management of distinct populations, especially with regard to the implications of modern climate change on southern relictual populations for this northern associated species.

References

Aubry, K.B., Statham, M.J., Sacks, B.N., Perrine, J.D., & Wisely, S.M. (2009) Phylogeography of the North American red fox: Vicariance in Pleistocene forest refugia. *Molecular Ecology*, **18**, 2668–2686.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C. a, & Saunders, N.C. (1987) Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Bidlack, A.L. & Cook, J.A. (2001) Reduced genetic variation in insular northern flying squirrels along the North Pacific Coast (Bidlack and Cook 2001). *Animal Conservation*, **4**, 283–290.

Brunhoff, C., Galbreath, K.E., Fedorov, V.B., Cook, J.A., & Jaarola, M. (2003) Holarctic phylogeography of the root vole (Microtus oeconomus): Implications for late Quaternary biogeography of high latitudes. *Molecular Ecology*, **12**, 957–968.

Brunsfeld, S.J., Sullivan, J., Soltis, D.E., & Soltis, P.S. (2000) Chapter 15 Comparative phylogeography of northwestern North America : a synthesis. 319–340.

DeBry, R.W. (1992) Biogeography of New World Taiga-Dwelling Microtus (Mammalia : Arvicolidae): A Hypothesis Test That Accounts for Phylogenetic Uncertainty. Author(s): Ronald W. DeBry Published by: Society for the Study of Evolution Stable URL : http://www.jstor.org/s. *Evolution*, **46**, 1347–1357.

DeChaine, E.G. (2008) A bridge or a barrier? Beringia's influence on the distribution and diversity of tundra plants. *Plant Ecology & Diversity*, **1**, 197–207.

Dragoo, J.W., Lackey, J.A., Moore, K.E., Lessa, E.P., Cook, J.A., & Yates, T.L. (2006) Phylogeography of the deer mouse (*Peromyscus maniculatus*) provides a predictive framework for research on hantaviruses. *Journal of General Virology*, **87**, 1997–2003.

Fennessy, J., Bidon, T., Reuss, F., Vamberger, M., Fritz, U., Janke, A., Kumar V., Elkan P., Nilsson, M.A., & Janke, A. (2016) Multi-locus Analyses Reveal Four Giraffe Species Instead of One. *Current Biology*, **26**, 1–7.

Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.

Hofreiter, M. & Stewart, J. (2009) Ecological Change , Range Fluctuations and Population Dynamics during the Pleistocene Review. *Current Biology*, **19**, R584–R594.

Hope, A.G., Waltari, E., Dokuchaev, N.E., Abramov, S., Dupal, T., Tsvetkova, A., Henttonen, H., MacDonald, S.O., & Cook, J.A. (2010) High-latitude diversification within Eurasian least shrews and Alaska tiny shrews (Soricidae). *Journal of Mammalogy*, **91**, 1041–1057.

Hotaling, A.S., Percival, H.F., Kitchens, W.M., & Kasbohm, J.W. (2010) The persistence of endangered Florida Salt Marsh Voles in salt marshes of the Central Florida Gulf Coast. *Southeastern Naturalist*, **9**, 795–802.

Jaarola, M., Martínková, N., Gündüz, I., Brunhoff, C., Zima, J., Nadachowski, A., Amori,

G., Bulatova, N.S., Chondropoulos, B., Fraguedakis-Tsolis, S., González-Esteban, J., José López-Fuster, M., Kandaurov, A.S., Kefelioğlu, H., Da Luz Mathias, M., Villate, I., & Searle, J.B. (2004) Molecular phylogeny of the speciose vole genus Microtus (Arvicolinae, Rodentia) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **33**, 647–663.

Kohli, B.A., Fedorov, V.B., Waltari, E., & Cook, J.A. (2015) Phylogeography of a Holarctic rodent (*Myodes rutilus*): Testing high-latitude biogeographical hypotheses and the dynamics of range shifts. *Journal of Biogeography*, **42**, 377–389.

Malaney, J.L. & Cook, J.A. (2013) Using biogeographical history to inform conservation: The case of Preble's meadow jumping mouse. *Molecular Ecology*, **22**, 6000–6017.

Reeder, D.M., Helgan, K.M., & Wilson, D.E. (2007) Global trends and biases in new mammal species discoveries. *Occasional Papers Museum of Texas*, **269**, 1–35.

Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., & Soltis, P.S. (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Chapter 2

Conservation challenges in North America: the precarious future of peripheral populations (and species) of the widely distributed meadow voles

Abstract

Current taxonomy holds that the meadow vole (*Microtus pennsylvanicus*) is the most widely distributed vole in North America composed of 28 subspecies that were described largely based on morphological characteristics. We used the first multilocus approach with range-wide sampling to characterize the genetic variation found within *M. pennsylvanicus*. We sampled 20 of 28 subspecies and amplified the mitochondrial gene Cytochrome b (Cytb) to explore genetic structure within this species. Lineage distance was put into the context of sister species divergence by constructing a comprehensive *Microtus* phylogeny and evaluating highly supported sister species relationships. We found strong support for four highly divergent clades within *M. pennsylvanicus*, the intraspecific clade distance was compared to interspecific sister species distance. Our findings revealed that intraspecific clade divergence overlapped with that seen between *Microtus* sister species. Finally, we used Bayes Factor species Delimitation (BFD) analyses to statistically test for the presence of multiple species within *M. pennsylvanicus.* BFD consistently and decisively supported multiple species based on mitochondrial and nuclear data. We suggest that our two westernmost clades be lumped to form *M. drummondii,* the eastern clade retain the name *M. pennsylvanicus*, and the coastal Florida populations be elevated to *M. dukecampbelli*. Finally, our findings highlight the importance that taxonomic strategies aim to conserve southern peripheral

populations, as they often disproportionately harbor high levels of genetic variation for widely distributed organisms.

Donavan J. Jackson [a], and Joseph A. Cook[a*]

[a]Department of Biology and Museum of Southwestern Biology. University of New Mexico, Albuquerque NM 87131-1051

[*]Co-authors for submission of a version of this chapter to a peer-reviewed journal

Introduction

A vast majority of intraspecific nomenclature for mammals is based on limited assessments of morphological characters (e.g., Hall 1981), which may be problematic given that a robust understanding of geographic variation, as reflected in taxonomy, is key to many aspects of an organism's biology including fundamental questions in ecology, evolution, biogeography, and zoonotic pathogen emergence (Dragoo *et al.* 2006). Taxonomy should reflect geographic variation and evolutionary relationships (Mayr 1942) and should be expected to occasionally change as systematists re-examine existing taxonomic hypotheses using new approaches and more extensive geographic sampling. Re-evaluations of well-known species recently uncovered surprising levels of cryptic variation and, in some cases, new species (Malaney *et al*. 2013; Fennessy *et al.* 2016). Phylogenetically-informed taxonomy is also necessary to establish robust conservation strategies (Mace 2004; Dawson *et al*. 2014), land management practices (Dawson *et al*. 2007), and ultimately compliance with laws, such as the Endangered Species Act (ESA). Implementation of the ESA has proven problematic when

management actions lacked a solid foundation for interpreting intraspecific geographic variation (e.g., Laerm *et al*. 1982; Avise 1994; Malaney *et al*. 2013). Molecular techniques are now routinely utilized for all major taxonomic groups, to test provisional taxonomic hypotheses, identify distinct evolutionary lineages, and detail their spatial distribution.

With over 60 extant species currently recognized (Musser & Carleton 2005), *Microtus* (Schank 1798) is one of the most speciose and ecologically diverse clades of Holarctic rodents. These species often occupy grassland and meadow ecosystems, but are also commonly found in alpine, steppe, taiga, and tundra (Getz 1985; Hoffmann & Koeppl 1985) and constitute a considerable proportion of the small mammal biomass in the Northern Hemisphere (French 1978; Pruitt 1968). Both the fossil record (Gromov & Polyakov 1992) and phylogenetic analyses (Conroy & Cook 2000; Fink *et al*. 2010) support the hypothesis that all 20 species of *Microtus* endemic to the Nearctic can be traced to a single colonization event across the Bering Land Bridge in the middle Pleistocene ~2.1 mya (Repenning *et al.* 1990), followed by the more recent expansion of a single Holarctic species, the tundra vole (*M. oeconomus)*, into North America (Conroy & Cook 2000; Fink *et al*. 2010).

Evolutionary relationships of the 66 nominal species of *Microtus* have been enigmatic in part due to rapid pulses of speciation (Repenning *et al.* 1990; Conroy & Cook 2000; Jaarola *et al.* 2004) and high levels of morphological convergence within arvicolids (Courant *et al*. 1997). Investigations within the last few decades have allocated species to different genera (e.g., *Neodon irene*: Musser & Carleton 2005; *Lasiopodomys gregalis*: Petrova *et al*. 2014), synonymized multiple species (e.g., *M. rossiaemeridionalis* and *M.*

levis: Musser and Carleton 2005), and identified new species (e.g., *M. anatolicus*: Kryštufek & Kefelioğlu 2002; and *M. qazvinensis*: Golenishchev 2003). Molecular phylogenetic analyses among species (Conroy & Cook 2000; Jaarola *et al*. 2004; Bannikova *et al.* 2010; Martinkova *et al*. 2012) and phylogeographic analyses within species (Brunoff *et al.* 2003; Conroy & Neuwald 2008; Tougard *et al*. 2013) have begun to shed light on the biogeographic and evolutionary histories of these voles, elucidating relationships and clarifying species limits not evident based on morphology alone (Jaarola & Searle 2002; Hellborg 2005).

We focus on the meadow vole, *Microtus pennsylvanicus* (Ord 1815), a Nearctic rodent primarily inhabiting grassy meadows, fields, and marshes associated with North American forests (Getz 1985). Meadow voles have one of the widest ranges among North American mammals, extending from interior Alaska, throughout Canada to Labrador, and south along the Rocky Mountains into northern New Mexico, as well as along the eastern seaboard of the Appalachian Mountains into northern Georgia and Florida (Figure 1). Insular populations are located along the Pacific Coast within the Alexander Archipelago of southeast Alaska and along the Atlantic Coast on several islands near New England and eastern Canada (Hoffmann and Koeppl 1985). Fossils indicate that the meadow vole was the most widely distributed microtine during the Pleistocene with a range extending south of its contemporary distribution into Texas, Louisiana, and southern New Mexico (Martin 1968; Smartt 1977; Kurten and Anderson 1980). To contextualize infraspecific variation in meadow voles, we first explore species limits as defined by the broader mitochondrial DNA (mtDNA) phylogeny of *Microtus* (Martinkova *et al.* 2007, Krystufek *et al.* 2012, Mahmoudi *et al.* 2014).

Hoffmann & Koeppl (1985) recognized 28 subspecies of *M. pennsylvanicus* (Figure 2). Another closely related insular species, *M. breweri*, was previously considered a subspecies but was later elevated to species based on morphological characteristics (Moyer *et al*. 1988). Six of subspecies of meadow voles have been listed as potentially threatened based on their limited spatial extent (MacDonald *et al.* 1998), two others are now presumed extinct (Hoffmann & Koeppl 1985; List *et al*. 2010), and another is federally protected under the Endangered Species Act (Appendix B). Historically, subspecific classifications for voles were based on univariate analyses of relatively few morphological characteristics (e.g., Youngman 1967), but morphometric studies increasingly use multivariate analyses to examine geographic variation (e.g., Moyer *et al*. 1988; Lowry 1998). Snell and Cunnison (1983) completed the most comprehensive multivariate assessment of morphological variation of meadow voles and concluded that skull measurements in *M. pennsylvanicus* generally failed to identify groups based on geographic proximity.

Molecular techniques have not been applied widely to *M. pennsylvanicus* and without comprehensive spatial sampling, circumscribing subspecies can be problematic (Conroy and Neuwald 2008; Malaney *et al*. 2013). Our primary focus is to explore spatial patterns of molecular variation throughout the vast range of the meadow vole and its close relative, *M. breweri*, to examine whether genetic structure mirrors taxonomic designations based on morphology (Hall 1981). We hypothesize, based on phylogeographic studies of other widespread northern mammals (Montuire *et al.* 2013; Hope *et al.* 2014; Kohli *et al.* 2014), that multiple divergent lineages will be identified within *M. pennsylvanicus* due to prolonged isolation in distinct geographic regions of

North America.

Materials and Methods

Analyses were carried out in 3 phases. Initially we assessed genetic divergence between sister species in *Microtus*. We then focused on mtDNA geographic structure within meadow voles and subsequently followed that with multilocus analyses to test the taxonomic validity of the independent mtDNA lineages identified in *M. pennsylvanicus* The first phase of analyses utilized Cytb sequences from all species in the genus *Microtus* to place the level of mtDNA divergence within *M. pennsylvanicus* in the context of interspecific divergence across the genus. Then, we used mitochondrial Cytb sequences from *M. pennsylvanicus* individuals to delimit distinct lineages and their geographic extent. The final phase incorporated mitochondrial and 6 independent nuclear loci to test five hypotheses related to geographic variation using a multilocus, coalescent-based species' delimitation method. The first hypothesis, H_1, is that *M. pennsylvanicus* is a monophyletic species (Carleton & Musser 2005). In our second hypothesis, H_{2a}, we tested the hypothesis of a major east/west dichotomy in *M. pennsylvanicus*, potentially identifying two incipient species, that is based on restricted fragment length polymorphisms (Plante *et al.* 1989). This split corresponds to a geographic break in lineages of other widespread mammals in North America (Reding *et al*. 2012). Additionally, we tested a second scenario of two species within *M. pennsylvanicus*. In this hypothesis, H_{2b}, all northern populations of *M. pennsylvanicus* were grouped as a single lineage with the Florida population being independent. The Atlantic and Gulf Coast constitute a break in lineages for several taxa in the Southeast United States (Avise

et al. 1987; Soltis *et al.* 2006). We then tested our fourth hypothesis, H_3, whether there are three incipient species in *M. pennsylvanicus*: a western group, eastern group, and Florida group. Finally, H_4, we tested for the presence of more than 3 independent lineages, as previous studies have identified in other North American cricetids (Conroy & Cook 2000).

I) *Microtus* phylogeny inferred from Cytochrome b

Sampling and Phylogenetic Reconstruction

We analyzed interspecific Cytb variation between *Microtus* sister species, generally accepting the species designations of Carleton & Musser (2005). Cytb sequences were downloaded for 112 individuals of 63 species from GenBank with 10 new sequences (two of which included the first genetic data for *M. breweri;* Appendix B) ranging from 489-1140 bp in length. Exceptions to Carleton and Musser (2005) were *M. hartingi* (Krystufek *et al.* 2012), *M. gromovi* (Bannikova *et al.* 2010), *M. atticus* (Rovatsos & Giagia-Athanasopoulou 2012), and two controversial species *M. rossiaemeridionalis* and *M. obscurus* (Jaarola *et al.* 2004; Bannikova 2010; Tougard *et al*. 2013; Markova *et al.* 2014). We treat the latter two taxa as independent species. In addition, we included one randomly chosen individual from each independent lineage in *M. pennsylvanicus.* All sequences were aligned in Geneious using the MUSCLE (Edgar 2004) algorithm and phylogenetic reconstruction was carried out in BEAST using the best substitution model, GTR+I+G, from jModelTest with a strict clock. We used a Coalescent Constant Population tree prior that was run for 40 million generations and sampled every 2,000 generations with a 20% burn in.

Genetic Distance

Genetic divergence of Cytb sequences between all sister species (excluding *M. pennsylvanicus* and its sister species *M. montanus*) were computed using the program MEGA7 (Kumar *et al.* 2016). All individuals were grouped by species if more than one individual was used in analyses and sister relationships were inferred from our Bayesian phylogeny (Figure 4). Mean between-group distance was computed under the Kimura 2-parameter model and standard errors (*SE*) were bootstrapped with 1,000 replicates, with pairwise deletions for missing data. Furthermore, we compared all *M. pennsylvanicus* Cytb sequences (50 individuals in 4 distinct monophyletic clades). For each clade, we used the same settings and calculated both the intraclade and interclade (clades within *M. pennsylvanicus*) mean distances. Finally, interclade mean distance was compared to the distance between highly supported sister species of *Microtus*. These comparisons allowed us to place divergence levels between highly supported clades in *M. pennsylvanicus* within the context of divergence between sister species in *Microtus*.

II) mtDNA variation and structure within *Microtus pennsylvanicus*

Sampling

We sampled populations spanning the geographic range of *M. pennsylvanicus* targeting 1-2 individuals of each subspecies as available. Sampling totaled 50 specimens (7 study skin clips, 3 muscle subsamples preserved in ethanol, 40 frozen tissue samples) from 43 localities (20 of the 28 named subspecies; Hoffmann & Koeppl 1985). Specimen loans were acquired from the Museum of Southwestern Biology (MSB), Florida Museum of Natural History (FLMNH), Museum of Vertebrate Zoology (MVZ), North Carolina Museum of Natural Sciences (NCSM), University of Alaska Museum of the North (UAM), University of Michigan Museum of Zoology (UMMZ) and University of

Washington Burke Museum (UWBM). Lastly, 5 closely related species (Conroy & Cook 2000) were chosen as outgroups for phylogenetic analyses: *M. breweri, M. canicaudus*, *M. longicaudus, M. montanus, and M. townsendii*.

PCR Amplification and Sequencing

Samples preserved in ethanol were washed overnight using STE buffer before extraction. Genomic DNA was extracted from skin clips, liver, or muscle samples using a standard salt extraction or QIAamp DNA Mini Kit (Qiagen Inc. Valencia, California). We amplified the full length of Cytb (1,140 bp) using the primer pair MSB05/MSB14 (Hope *et al.* 2010) and PCR reactions that consisted of: 1-2 µL DNA template, 2.5 µL 10x PCR buffer, 2 µL 25 mM $MgCl_2$, 1 µL bovine serum albumin, 0.5 µL deoxynucleoside triphosphates, 0.13 µL AmpliTaq DNA Polymerase (Applied Biosystems, Foster City, California), 1 µL of each primer and remaining volume of H_2O to total 25 µL. All PCRs were performed in a T100 thermocycler (Bio-Rad, Hercules, California) with initial denaturation at 95°C for 2 minutes and 40 cycles of denaturation 95°C for 15 seconds, annealing for 30 seconds (51°C), and extension at 72°C for 1 minute per kb (variable) with a final extension at 72°C for 5 minutes preceding cooling at 4°C. PCR products were visualized by agarose gel electrophoresis and cleaned with ExoSap-It (Affymatrix, Santa Clara, California) or 30% polyethylene glycol precipitation. Cleaned PCR products were cycle sequenced using the ABI BigDye version 3.1 sequencing kit (Applied Biosystems, Foster City, California) following the conditions of Platt *et al*. (2007). Cycle sequencing products were cleaned and precipitated following the sodium acetate, ethylenediaminetetraacetic acid (EDTA) and ethanol

protocol. Automated sequencing was conducted at the University of New Mexico, Molecular Biology Facility using an ABI 3100 DNA sequencer.

Sequences were edited and aligned using the MUSCLE algorithm (Edgar 2004) in Geneious v8.0.5 and visually confirmed. We established that no nuclear copies of mitochondrial DNA were sequenced, a known issue for Cytb in Arvicolinae (DeWoody *et al*. 1999), by translating nucleotide bases to amino acids and confirming no stop codons or unexpected patterns of replacement were present at any codon positions.

Phylogenetic Reconstruction

Our sampling spans the geographic distribution of *M. pennsylvanicus*, allowing exploration of geographic structure and tests of the morphology-based subspecific designations. Phylogenetic reconstruction of *M. pennsylvanicus* was conducted using Bayesian inference and the Markov chain Monte Carlo (MCMC) search implemented in BEAST v1.8.2 (Drummond *et al.* 2012). Prior to MCMC, the best substitution model was determined to be TrN+G in jModelTest v.2.1.7 (Darriba *et al.* 2012). Two runs were employed with identical sampling parameters of 40 million generations, sampling every 2000 generations. We used a mutation rate of 0.0772 (Hope *et al.* 2014) with a strict clock and coalescent constant population tree prior. Resulting tree files were then summarized in TreeAnnotator v1.8.2 (using a 20% burn in) and assessed in TRACER v1.6.0 (Rambaut & Drummond 2007) ensuring Effective Sample Size (ESS) was above 200 for all parameters and that independent runs had converged.

III) Species tree estimation and Bayes Factor species Delimitation in *M. pennsylvanicus*

For species tree analyses, we used a subset of 30 samples sequenced at 7 independent loci, mitochondrial Cytb and six nuclear loci: Protein C-est-2 (ETS2: 875 bp), ß-fibrinogen (FGB: 600 bp), Growth Hormone Receptor exon 10 (GHR: 460 bp), the first exon of Interphotorecepter Retinoid-Binding Protein (IRBP: 625 bp), Lecithin: Cholesterol acyltransferase (LCAT: 465 bp) and Recombination Activating Protein 1 (RAG1: 860 bp). Primer sequences used for analyses are listed with corresponding references in Appendix D. Samples to be sequenced for these 7 loci were chosen with the aim of incorporating individuals that were spread throughout the geographic range within each mitochondrial (mtDNA) clade.

To infer alleles of nuclear heterozygotes we used Phase v2.1.1 (Stephens *et al*. 2001; Stephens & Donnelly 2003), which implements a Bayesian statistical method to reconstruct haplotypes for nuclear sequences that include multiple bases that are heterozygous within an individual. Heterozygotes were inferred using a burn-in of 10,000 iterations and a run length of 10,000 iterations (McCormack *et al*. 2011) resulting in two phased nuclear haplotypes. For all analyses we used one randomly chosen haplotype for each individual. Nucleotides were only retained if their posterior probabilities were >0.90, the remaining ambiguous sites were recoded as "N" to prevent bias in phylogenetic analyses.

A Bayes Factor species Delimitation (BFD) approach was used to compare alternative species trees from *BEAST (Grummer *et al*. 2014) and determine the number of diagnosable evolutionary lineages in *M. pennsylvanicus*. Bayes Factors (2*ln*Bf) were used to determine the support of the most likely scenario given the data, as these methods allow for flexibility in testing various hypotheses for a given system (Grummer *et al.*

2014). These species delimitation analyses have four steps: (1) each individual is assigned to a lineage (or "species") based on previous studies or exploratory analyses; (2) alternative groupings are created by combining, splitting, or reassigning individuals to lineages; (3) a species tree is generated and the marginal likelihood estimate (MLE) is calculated for each model or hypothesis; and (4) the Bayes factors for each hypothesis are determined, yielding the best species delimitation model given the data (Grummer *et al.* 2014). Following the completion of all BFD models, all hypotheses were scored relative to the best hypothesis (the model with largest MLE). All 2*ln*Bf were scored with competing models using the recommendations of Kass & Raftery (1995) that provide assessments of strength in support of a given model: a 2*ln*Bf of 0-2 means "not worth more than a bare mention", 2*ln*Bf of 2-6 means "positive" support, 2*ln*Bf of 6-10 provides "strong" support and a 2*ln*Bf >10 is indicative of "decisive" support between the two models being compared.

We performed BFD to test the null hypothesis that *M. pennsylvanicus* is a single monophyletic species (current taxonomy is correct) against alternative hypotheses that differentially grouped the mitochondrial clades: Hypothesis H_1: single species; Hypothesis H_{2a}: two species largely split East/West, as suggested based on restricted fragment length polymorphism assays (Plante *et al.* 1987); Hypothesis H_{2b}: a second two species model that grouped the 3 wide-ranging clades and treated the Florida clade as an independent lineage. The second clade (Florida) is based on distinctive morphological characteristics for the Florida populations (Woods *et al.* 1982) and the repeated Atlantic Coast vs. Gulf Coast phylogeographic break shown in >25 other organisms (Soltis *et al*. 2006); Hypothesis H_3: three species, Western (composed of the Northwestern and Central

clades), Eastern, and Florida; and our final hypothesis H_4 treats each Cytb lineage as an independent species.

The five competing scenarios of species delimitation models were run under the multispecies coalescent model implemented in *BEAST (Heled & Drummond 2010) in BEAST v.1.8.2 (Drummond *et al.* 2010). For each model tested, we generated a separate xml file in BEAUti with individuals assigned specifically for each hypothesis. Site heterogeneity models were assigned based on jModelTest (Darriba *et al.* 2012) and are shown in Appendix D. A mutation rate of 0.0772 was assigned to Cytb (Hope *et al.* 2014) and all nuclear loci were estimated relative to Cytb. We used a Yule Process species tree prior and proper ploidy was assigned for all loci. We conducted two independent runs for each hypothesis for 250 million generations sampling trees every 10,000 generations. Marginal likelihood estimates (MLE) were calculated using both path sampling (PS, Xie *et al.* 2011) and stepping-stone analyses (SS, Baele *et al.* 2012). Following each BEAST run, we ran PS and SS analyses for 250 steps and each step for 1 million generations (for a total of 250 million generations). The species tree for each model was assessed in Tracer, ensuring all ESS values were over 200. All analyses were carried out in two separate groupings. The first used both mtDNA and nuDNA and to explore the influence of the mtDNA locus, we also ran analyses with just the six independent nuDNA loci.

Results

I) *Microtus* phylogeny inferred from Cytochrome b

The phylogeny for *Microtus* was constructed with Cytb sequences (n = 122). This is the most comprehensive phylogeny for *Microtus* incorporating 63 presumed

species. We included two recently proposed species (*M. gromovi*-Bannikova *et al.* 2010; *M. atticus* -Rovatsos & Giagia-Athanasopoulou 2012), uncertain species (*M. rossiaemeridionalis and M. obscurus* – Musser & Carleton 2005), and provided the first mitochondrial sequences for *M. breweri*. Relationships across the phylogeny are largely congruent with previous studies (Conroy & Cook 2000; Jaarola *et al.* 2004; Martinkova *et al.* 2012). Mitochondrial data nest *M. breweri* within *M. pennsylvanicus* (Figure 4). Across this phylogeny, genetic distances between 13 sister species pairs averaged 5% (range 1.5-9.7%). Lowest divergence values were between *M. abbreviatus – M. miurus* (1.5%), *M. bavaricus – M. liechtensteini* (1.6%), and *M. breweri – M. pennsylvanicus* (1.8%) while highest sister species divergence was between *M. daghestanicus –M. subterraneus* (9.7%).

II) Genetic variation and structure within *Microtus pennsylvanicus*

Cytochrome b gene tree analyses incorporated *M. pennsylvanicus* specimens (n = 50) that were broadly distributed throughout North America. No pseudogenes were observed and bias in variation across codons was typical of functional Cytb (Irwin *et al.* 1991). Bayesian species tree reconstruction strongly supported 4 geographically structured mtDNA clades: Northwestern, Central, Eastern, and Florida (Figure 1). Northwestern includes populations from Alaska, northern Canadian provinces, and British Columbia. Central encompasses Saskatchewan and Manitoba and extends south along the Rockies into northern New Mexico and the recently extirpated Chihuahuan meadow vole (List *et al.* 2010). Eastern spans northeastern Canada and extends south along the Appalachian Mountains to Georgia. Finally, the samples from Florida represent a distinctive, highly restricted population located in a salt marsh along the Gulf Coast of

Florida. Northwestern, Central, and Eastern clades are reciprocally monophyletic; however, relationships between these clades remain uncertain due to low support.

Molecular analyses generally did not correspond to subspecific designations. Notably, two widely distributed subspecies are each split between two distinct mtDNA clades, *M. p. drummondii* in the Central and Northwestern clades and *M. p. pennsylvanicus* in the Central and Eastern clades (Figure 3). Additionally, we rarely recovered sister relationships for multiple individuals from the same subspecies (Figure 2). Within each clade, highly supported groups will require more comprehensive sampling to better define the geographic extent of clade substructure. For example, within the Northwestern clade individuals on Admiralty Island are distinctive (consistent with *M. p. admiraltiae*) as well as on Kadin Island and the nearby mainland (*M. p. rubidus*) in Southeast Alaska. In the Central clade, two specimens representing *M. p. modestus* from the Sangre de Cristo Mountains of northern New Mexico also are highly divergent (average genetic difference of 1.7%) from nearby populations in other mountain ranges in New Mexico, Colorado and Mexico. Additionally, *M. breweri* does not appear to be distinct as it was nested within the Eastern clade of *M. pennsylvanicus*. Average pairwise genetic difference between the 4 Cytb clades (Central, Eastern, Florida and Northwestern) of *M. pennsylvanicus* is 3.9% (Table 1). Lowest divergence is between the Northwestern and Central clades (1.8%) which is similar to the genetic distance between *M. breweri* and the Eastern clade. The greatest divergence is between the Florida and Central clades (4.6%). All 3 individuals from the Florida population shared a single haplotype. Pairwise difference analyses between all *M. pennsylvanicus* individuals and *M. montanus* ranged from 6.6-7.8% (0.001-0.008 *SE*).

III) Bayes Factor Species Delimitation in *M. pennsylvanicus*

All BFD analyses decisively supported a scenario in which *M. pennsylvanicus* is composed of multiple species. For our first set of analyses that included mtDNA and nuDNA marginal likelihood estimations were nearly identical between the PS and SS analyses and the order in which the 4 species distribution hypotheses received support was consistently the same between the two approaches (Table 2). Based on BFD, the PS and SS estimations repeatedly supported H_4 (4 major lineages as 4 distinct species) as the best systematic explanation for variation within *M. pennsylvanicus*. In this scenario, our 4 species model received "decisive" support ($2ln$Bf = 16-105.28 for PS and 15.58-105.14 for SS) over alternative scenarios. All other models exhibited progressively less support as the number of distinct evolutionary lineages was reduced. Alternatively, when only nuclear loci were used in BFD analyses, support for the 4 species model (H_4) was no longer decisively supported. Conversely, the 2 species model (H_{2a}) and 3 species model (H_3) received nearly identical PS and SS support (Table 2) as the best systematic explanation.

Discussion

Key aspects relating to the taxonomy of species in the genus *Microtus* remain uncertain due to high levels of morphological convergence (Courant *et al.* 1997); however, molecular analyses are beginning to elucidate evolutionary relationships, document previously cryptic species, and explore species limits. High levels of taxonomic fluidity may be partially attributed to better resolution of the history of diversification through molecular studies that include more extensive geographic

sampling and often reveal high levels of intraspecific divergence (Jaarola *et al.* 2004; Bannikova *et al.* 2010, Martinkova & Moravec 2012*).* We found that mitochondrial data strongly supported four monophyletic clades that are geographically distributed. One of these has an extremely restricted range in Florida, while the other three are more widely distributed. Additionally, genetic divergence between each clade falls well within the range that is supported between sister species in *Microtus*, emphasizing that these highly divergent clades may be on the cusp of incipient speciation. When these clades were tested with multilocus data, BFD analyses strongly supported the presence of three species within *M. pennsylvanicus*. With regard to conservation, this molecular assessment of *M. pennsylvanicus* provides a more robust geographic framework for understanding how southern peripheral populations contribute to genetic variation harbored by this wide spread species.

Genetic Structure within Microtus pennsylvanicus

Discrepancies between morphological subspecies and molecular clades have been routinely noted in other species of *Microtus* (Conroy & Cook 2000; Jaarola & Searle 2002; Brunhoff *et al.* 2003; Conroy & Neuwald 2008; Sawyer & Cook 2016). In *M. pennsylvanicus,* geographic distribution of Cytb clades generally did not coincide with subspecies boundaries (Figure 2). Two subspecies in SE Alaska, *M. p. admiraltae* and *M. p. rubidus*, were genetically distinctive, potentially reflecting the dynamic glacial history of the region where several other mammalian species persisted and diverged through one or more of the Late Pleistocene glacial advances (e.g., Fleming & Cook 2002). In the Central clade, a population representing *M. p. modestus* from the Sangre de Cristo Mountains of New Mexico at the southern tail of the Rocky Mountains, was significantly

divergent. Expanded analyses that integrate more extensive geographic sampling of these three distinctive subspecies, and integrates multivariate analyses of morphological characteristics, are needed to fully explore and delineate the contribution of these populations to overall genetic diversity. The distinctive population in northern New Mexico raises the issue of whether other southern populations south of the Rockies (e.g., in the Gila region of New Mexico) were also distinct. Those peripheral populations have apparently been extirpated in the past 50 years (Hubbard *et al.* 1983; List *et al.* 2010; Jones 2016) and may constitute a loss of significant genetic diversity within the species (List *et al*. 2010).

This multilocus molecular analysis revealed 3 monophyletic clades throughout the continental range of *M. pennsylvanicus*. Previous studies focused on arvicolines have shown that isolated populations can diverge quickly (e.g. Martinkova *et al*. 2007; Krystufek *et al.* 2009; Sawyer & Cook 2016). Rapid divergence highlights the value of a fuller understanding of how cyclical glacial events of the Pleistocene played a role in structuring contemporary genetic diversity within mammals. Divergent Florida samples represent a relict population now restricted to salt marsh habitat along the Florida Gulf Coast. McCleery & Zweig (2016) reported capturing individuals from this population at an additional 8 sites along 85 km of Florida coastline; however, the specific localities were not reported, leaving the geographic extent of this lineage unclear. Farther north, the Eastern clade occupies much of Canada and the United States east of Ontario and Ohio and north of South Carolina. Geographic disjunction between the Eastern and Central clades was previously identified in *M. pennsylvanicus* using mtDNA restriction fragment analyses (Plante *et al.* 1987). This apparent zone of contact is shared with other widely

distributed vertebrates such as black bear (*Ursus americanus*; Puckett *et al.* 2015), bobcat (*Lynx rufus*; Reding *et al.* 2012), and snakes (e.g., *Coluber constrictor*; Burbrink *et al.* 2007). Often termed suture zones, these areas of contact that are shared between divergent lineages of many species likely reflect the common influence of historical barriers or similar patterns of recolonization following climate shifts. Furthermore, the region of contact between the Northwestern and Central clades falls in the proximity of the Rocky Mountains along the border of British Columbia and Alberta. This region has been characterized as a suture zone for many western North American organisms (Swenson and Howard 2005). Similarly, the Atlantic Coast-Gulf Coast phylogeographic break reported for other organisms (e.g., Avise *et al.* 1987; Soltis *et al*. 2006) coincides with the relatively high mtDNA divergence (4.1%; Table 1) between the Florida and Eastern clades of meadow voles.

Microtus *Systematics and Taxonomic Recommendations*

Our goal was to sufficiently characterize infraspecific variation within *M. pennsylvanicus* to assess whether the taxonomy should be revised. Intraspecific Cytb mitochondrial distances generally average around 2% within species of *Microtus* (Jaarola *et al.* 2004; Baker & Bradley 2006), but within *M. pennsylvanicus*, average mtDNA distance between the 4 monophyletic clades is 3.9%. The Holarctic vole, *M. oeconomus,* another widespread microtine, also exhibited comparable levels of intraspecific divergence ranging from 2.7%-4.3% (Brunhoff *et al.* 2003). While further analyses may strengthen our understanding of species' limits*,* our data shows that 4 divergent lineages within *M. pennsylvanicus* may be on the cusp of incipient speciation when based on interclade comparisons.

Adding to the complexity of taxonomy in this genus, interspecific mtDNA distances between sister species in *Microtus* range between 1.5%-9.7% (Appendix A), overlapping values for intraspecific divergence. The closest related nominal pairwise comparison for species, *M. miurus* and *M. abbreviatus,* is only 1.5% divergent, but the validity of the later insular species has been questioned (Conroy & Cook 2000; MacDonald & Cook 2009; Weksler *et al*. 2010). Another species pair exhibiting low divergence (1.6%) is *M. bavaricus* and *M. liechtensteini*. For 6 other nominal (Musser & Carleton 2005; with the exception of *M. obscurus*) sister species pairs, estimates of genetic divergence are comparable or lower than the highly divergent clades in *M. pennsylvanicus*. Additionally, total divergence between *M. thomasi* and *M. atticus* is 2.9%; however, reproductive isolation has been observed (Rovatsos & Giagia-Athanasopoulou 2012). Recent phylogeographic studies have identified distances between putative species based on Cytb variation for *M. agrestis* (about 6%, divergent; Jaarola and Searle 2002)*, M. californicus* (4.5%; Conroy & Neuwald 2008)*, M. mexicanus* (8.2%; Crawford *et al.* 2011)*,* and *M. savii* (7.6%; Tougard *et al.* 2013). With regard to species identification, estimates of genetic distance in *Microtus* do not consistently reflect speciation, but infraspecific mtDNA variation between clades in *M. pennsylvanicus* is within the range of several species pairs previously examined.

Under the framework of the general concept of species (de Queiroz 2007), we agree with Bagley *et al.* (2016) that genealogical and statistical evidence from genetic data is sufficient for species identification. Multilocus DNA analyses consistently identify 4 independently evolving lineages within *M. pennsylvanicus* that we feel should be recognized as separate species, of these, the northwestern and central clades overlap in

southern British Columbia and the dynamics and consequences of contact should be carefully evaluated. We proceeded with species delimitation under the general lineage concept of species proposed by de Queiroz (2007), where we treat separately evolving metapopulation lineages as evidence for species designation. In our study, BFD estimates of species relationships within *M. pennsylvanicus* support the hypothesis of 4 independently evolving lineages (=species) based on statistical assessment of genetic evidence that incorporated one mtDNA gene and 6 nuDNA loci. Alternatively, using only nuDNA in BFD analyses failed to decisively support either the 2 species (split West and East) or a 3 species (West, East and Florida) model as the most likely scenario given the data. Multilocus molecular data consistently and decisively supported evolutionary models that comprised more than a single species (Table 2). In both the PS and SS estimates, current taxonomy was decisively shown to be the least supported scenario for species relationships. Our results suggest that all mitochondrial clades should be elevated to the species level based on both mtDNA and nuDNA. Mitochondrial data has been shown to be a reliable indicator of population history and species limits due to rapidly evolving DNA (Zink & Barrowclough 2008). However, the rapid evolution of mtDNA may drive species delimitation analyses. For these reasons, we advocate the more conservative approach of basing our taxonomic recommendations on BFD analyses using nuDNA loci only. The two species model (H_{2a}) received 2*ln*Bf support of only 1.4 and 1.26 (PS and SS respectively) over our 3 species model (H_3), which is barely worth mentioning (Kass & Raftery 1995). We suggest taxonomy should reflect model H_3, in which our Northwestern and Central clades are lumped to form a Western species, and the Eastern & Florida clades are treated as independent species. We suggest the Florida

clade be elevated based on the corroboration between distinct morphological characteristics that are comparable that observed in *M. breweri* (Woods 1992) and high genetic differentiation from other meadow vole populations (this study). Accordingly, we propose that the isolated Florida population be elevated to *M. dukecampbelli*. The Eastern clade should remain *M. pennsylvanicus* (Ord 1815), consistent with the type locality in Pennsylvania. The Central and Northwestern clade elevated to *M. drummondii* (Audubon & Bachman, 1854) as this is the oldest western subspecies. Suggested subspecies within each species is shown in Table 3.

Peripheral Populations of Mammals Under Global Warming Scenarios

Peripheral populations often harbor genetic signatures distinct from core populations and this variation may accumulate in a relatively short time and even play a potential role in peripatric or allopatric speciation (Barraclouch & Volgler 2000; Krystufek *et al.* 2009; Ikeda *et al.* 2012; Fernandez 2012; Castellanos-Morales *et al.* 2016). Isolated populations are often characterized by distinctive morphological, behavior, or life history characteristics (Fisler 1961; Anderson & Hubbard 1971; Snyder & Peterson 1999), emphasizing the value of assessing variation throughout the range of species. Often, isolates are sources of genetic diversity with high conservation value (Malaney & Cook 2013), so loss of peripheral populations harboring suites of rare characteristics and distinct genetic variation may decrease resilience of species that are facing rapidly changing environmental conditions (Frankel & Soule 1981; Gilpin & Soule 1986). Ehrlich (1988) emphasized the consequences of the loss of genetically distinctive populations.

This refined view of geographic variation within *M. pennsylvanicus* provides a foundation for establishing management priorities, especially with regard to conservation of peripheral populations. Peripheral populations generally are at greater risk of extinction due to multiple factors, including smaller range size, changing environments, and stochastic perturbations. Within *M. pennsylvanicus,* multiple isolates exist along the western, southern and eastern periphery of its range and several of these isolates face threats common to marginal populations (Lienert *et al.* 2002). To date, two marginal subspecies of *M. pennsylvanicus* have been reported as extinct (*M. p. nesophilus* along the eastern periphery and *M. p. chihuahuensis* in the Southwest). Two more locally isolated populations along the southern edge of the species distribution in central New Mexico also appear to have been extirpated (Anderson 1961; Hoffmann & Koeppl 1985; List *et al.* 2010). As meadow voles are strongly associated with the mesic habitats, distinctive populations along the southern distributional limits should be closely monitored as they persist in environments now undergoing rapid change (List *et al.* 2010; Woods *et al.* 1982; Hotaling *et al.* 2010). This pattern of tenuous southern populations is common to many other forest or mesic associated mammals, such as red squirrels (genus *Tamiasciurus*) and jumping mice (genus *Zapus*) where southern isolates maintain private alleles and distinctive morphology, behavior, and ecological adaptations in an environment experiencing significant habitat modification and loss (e.g., Koprowski *et al.* 2006; De Grammont & Cuaron 2008; Malaney & Cook 2013; Hope *et al.* 2016).

The Chihuahuan meadow vole (*M. p. chihuahuensis*), a Pleistocene relict, became restricted to a small marsh surrounded by an arid landscape in northern Chihuahua near Galeana, Mexico. Designated as a subspecies based on morphology (Bradley & Cockrum

1968), later mtDNA investigation found that they were approximately 0.07% divergent from *M. p. modestus* and 2.6% divergent from *M. p. pennsylvanicus* based on 379 bp of Cytb (List *et al.* 2010). Expanding upon this, we added a full sequence of Cytb (1140 bp) and found that *M. p. chihuahuensis* is 1.3% divergent from nearby New Mexico and Colorado populations. List *et al.* (2010) concluded that the likely cause of the extinction of this subspecies was the diversion of water for irrigation purposes from the marshes, resulting in the loss of the wetlands. Other mesic associated species in the arid Southwest have experienced similar fates including pupfishes (*Cyprinodon* sp.; Minckley & Marsh 2009) and the Meadow jumping mouse (*Zapus hudsonius luteus*; Frey & Malaney 2009). Absence of conservation efforts or protection resulted in the extinction of the endemic *M. p. chihuahuensis* from the reportedly suboptimal environment it survived in since the Last Glacial Maximum in northern Mexico (Bradley & Cockrum 1968). Farther north in New Mexico a similar case with two local extirpations of meadow voles and their habitat occurred, due primarily to climate driven loss of habitat and overgrazing of habitat through the introduction of domestic livestock (Anderson & Hubbard 1971), highlighting a major range contraction (approximately 700 km) of this species in western North America. Previously occupied localities in west-central New Mexico were surveyed throughout the 1970's (Hubbard *et al.* 1983) with no success. In 2015 we re-surveyed these sites, but again found no evidence of *M. pennsylvanicus*, but did capture *M. mogollonensis*. The recent extirpation of peripheral populations such as these highlights the role of effective management and protection against anthropogenic change for geographically restricted populations.

Herein, we present the first study that has investigated genetic divergence of the peripheral Florida salt marsh vole, *M. p. dukecampbelli,* using a multilocus (mtDNA and 6 nuDNA loci) BFD approach. An apparent disjunct relict, *M. p. dukecampbelli* is distinct both morphologically (Woods *et al.* 1982, Woods 1992) and based on mtDNA and nuDNA variation. Relatively few Florida salt marsh voles have been captured since 1979 (Woods *et al.* 1982, Hotaling *et al.* 2010, McCleery & Zweig 2016) and little is known of the life history of this population due primarily to difficulty in studying them. Density has not been estimated for this population, but is likely lower (Woods *et al.* 1982; Hotaling *et al.* 2010; Austin *et al.* 2014) than 70 individuals per hectare that have been documented for meadow voles in the marshes of Virginia (Block & Rose 2005). Once wide-ranging, fossil data indicate that *M. pennsylvanicus* was not widely distributed throughout Florida after 5,000 yBP (Woods *et al*. 1982). Because of their extremely restricted geographic range (Figure 1), the Florida salt marsh vole is vulnerable (Woods *et al.* 1982; Park *et al.* 1988; Woods 1992) to an increase in catastrophic weather, such as severe tropical storms and hurricanes (Hotaling *et al.* 2010), and the projected 1-2 meter rise in sea levels from global warming (Intergovernmental Panel on Climate Change 2007). Relic populations of prairie vole (*M. ochrogaster*) along the Gulf coast of Texas and Louisiana early in the 1900s were reported as abundant at several localities but are now likely extirpated (Lowery 1974), serving as a reminder of the vulnerability of the Florida salt marsh vole. Our multi-locus coalescent-based species delimitation methods (BFD) as well as previous morphological data (Woods 1982) strongly support the conclusion that this independently evolving lineage warrants species designation.

Conclusion

The considerable multilocus phylogeographic variation uncovered in the meadow vole reflects multiple incipient species that are robustly supported as distinctive by BFD analyses. Moreover, four geographically structured lineages have an average mtDNA distance comparable to or greater than pairwise differences observed in several other sister species pairs of *Microtus*. Our study reveals a highly divergent population along the Florida Gulf Coast, which should be elevated to the species level as *M. dukecampbelli*. This population likely represents what is the least understood species in North America and perhaps the smallest range of any North American mammal, by its restriction to a single isolated salt marsh in Florida (McCleery & Zweig 2016). In addition to genetic divergence, it harbors distinctive morphological characteristics and given predicted environmental change due to changing climatic conditions, this species may soon go extinct. Additionally, we propose that the Northwestern and Central clades be lumped and elevated as the western species *M. drummondii.* Finally, we propose the Eastern clade retain the binomial *M. pennsylvanicus*. Many of these peripheral populations are relics of the Pleistocene, often surviving in suboptimal conditions and fragile environments. For the newly defined *M. drummondii,* the extinction of the subspecies in Mexico and populations in New Mexico serve as a stark reminder of the vulnerability of isolated populations and point to the need to more closely monitor southern peripheral populations of other boreal or forest associated species. Future studies of meadow voles should incorporate more comprehensive sampling of the Southern Rockies, Atlantic island populations and Southeast Alaska along with adjacent British Columbia to better understand diversity in these regions. Finally, the island endemic *M. breweri* appears to

be conspecific with *M. pennsylvanicus*, a tentative conclusion that warrants more comprehensive examination integrating nuDNA.

References

Anderson, S. (1961) A relict population of *Microtus pennsylvanicus* in southwestern New Mexico. *American Museum Novitates*, **2034**, 1-3.

Anderson, S. & Hubbard J.P. (1971) Notes on Geographic Variation of *Microtus pennsylvanicus* (Mammalia, Rodentia) in New Mexico and Chihuahua. *American Museum Novitates*, **2460**, 1-8.

Audubon, J.J. & Bachman, J. (1854) *The viviparous quadrupeds of North America*. V.G. Audubon, New York, **2**, 1-334.

Austin, J.D., Saarinen, E.V., Arias-Perez, A., McCleery, R.A., & Lyons, R.H. (2014) Twenty-one novel microsatellite loci for the endangered Florida salt marsh vole (*Microtus pennsylvanicus dukecampbelli*). *Conservation Genetics Resources*, **6**, 637-639.

Avise, J.C. (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall. New York, NY.

Avise, J.C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., & Saunders, N.C. (1987) Intraspecific phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., & Alekseyenko, A.V. (2012) Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*. **29,** 2157-2167.

Baker, R.J., & Bradley, R.D. (2006) Speciation in Mammals and the Genetic Species Concept. *Journal of Mammalogy*, **87**, 643-662.

Bannikova, A.A., Lebedev, V.S., Lissovsky, A.A., Matrosova, V., Abramson, N.I., Obolenskaya, E.V. & Tesakov, A.S. (2010) Molecular phylogeny and evolution of the Asian lineage of vole genus *Microtus* (Rodentia: Arvicolinae) inferred from mitochondrial cytochrome *b* sequence. *Biological Journal of the Linnean Society,* **99**, 595-613.

Barraclough, T.G., & Vogler, A.P. (2000) Detecting the Geographical Pattern of Speciation from Species-Level Phylogenies. *The American Naturalist*, **155**, 419-434.

Block, C.P., & Rose, R.K. (2005) Population dynamics of *Oryzomys palustris* and *Microtus pennsylvanicus* in Virginia tidal marshes. *Northeastern Naturalist*, **12**, 295-306.

Bradley, W.G. & Cockrum, E.L. (1968) A new subspecies of meadow vole (*Microtus pennsylvanicus*) from northwestern Chihuahua, Mexico. *American Museum Novitates,* **2325**, 1-7.

Brunhoff, C., Galbreath, K.E., Fedorov, V.B., Cook, J.A. & Jaarola, M. (2003) Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary biogeography of high latitudes. *Molecular Ecology,* **12**, 957-968.

Burbrink, F.T., Fontanella, F., Pyron, R.A., Guiher, T.J. & Jimenez, C. (2007) Phylogeography across a continent: The evolutionary and demographic history of the North America racer (Serpentes: Colubridae: *Coluber constrictor*). *Molecular Phylogenetics and Evolution,* **47**, 274-288.

Castellanos-Morales, Gamez, N., Castillo-Gamez, R.A., & Eguiarte, L.E. (2016) Peripatric speciation of an endemic species driven by Pleistocene climate change: The case of the Mexican prairie dog (*Cynomes mexicanus*). *Molecular Phylogenetics and Evolution*, **94**, 171-181.

Conroy, C.J. & Cook, J.A. 2000. Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Rodentia: Muridae). *Molecular Ecology,* **9**, 165-175.

Conroy C.J. & Neuwald, J.L. (2008) Phylogeographic study of the California vole, *Microtus californicus*. *Journal of Mammalogy,* **89**, 755-767.

Courant, F., David, B., Laurin, B. & Chaline, J. (1997) Quantification of cranial convergences in arvicolids (Rodentia). *Biological Journal of the Linnean Society*, **62**, 505-517.

Darriba D., Taboada, G.L., Doallo, R., & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

De Grammont, P.C., & Cuarón, A. (2008) Tamiasciurus mearnsi. The IUCN Red List of Threatened Species. doi: http://dx.doi.org/10.

Department of the Interior. U.S. Fish and Wildlife Service. Endangered and Threatened Wildlife and Plants; Endangered Status for the Florida Salt Marsh Vole. *Federal Register,* **56**, 1457-1459.

DeWoody, J.A., Chesser, R.K. & Baker, R.J. (1999) A Translocated Mitochondrial Cytochrome *b* Pseudogene in Voles (Rodentia: *Microtus*). *Journal of Molecular Evolution,* **48**, 380-382.

Dragoo, J. W., Lackey, J.A., Moore, K.E., Lessa, E.P. Cook, J.A., & Yates, T.L. (2006) Phylogeography of the deer mouse *(Peromyscus maniculatus)* provides a predictive framework for research on hantaviruses. *Journal of General Virology*, **87**, 1997.

Drummond A.J., Suchard M.A., Xie, D. & Rambaut A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* **32**,1792-1797.

Ehrlich, P.R. (1988) The loss of diversity: causes and consequences. *Biodiversity* (ed. By E.O. Wilson), pp 21-27. National Academy Press, Washington, D.C.

Excoffier, L., & Lischer, H. E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources,* **10**, 564-567.

Fennessy, J., Bidon, T., Reuss, F., Kumar, V., Elkan, P., Nilsson, M.A., Vamberger ,M., Fritz, U., & Janke ,A. (2016) Multi-locus Analyses Reveal Four Giraffe Species Instead of One. *Current Biology*, **26**, 2543-2549.

Fernandez, J.A. (2012) Phylogenetics and biogeography of the microendemic rodent *Xerospermophilus perotensis* (Perote ground squirrel) in the Oriental Basin of Mexico. *Journal of Mammalogy*, **93**, 1431-1439.

Fink, S., Fishcher, M.C., Excoffier, L., & Heckel, G. (2010) Genomic scans support repetitive continental colonization events during the rapid radiation of voles (Rodentia: *Microtus*): the utility of AFLPs versus mitochondrial and nuclear sequence markers. *Systematic Biology*, **59**, 548-572.

Fisler, G.F. (1961) Behavior of salt-march *Microtus* during winter high tides. *Journal of Mammalogy,* **42**, 37-43.

Fleming, M.A. and J.A. Cook. 2002. Phylogeography of endemic ermine (*Mustela ermine*) in southeast Alaska. Molecular Ecology 11, 795-807.

Frankel, O.H., & Soule, M.E. (1981) *Conservation and evolution*. Cambridge University Press, Cambridge, England.

French, N.R. (1978) Small mammals as components of the consumer system pages 61-67

in Snyder, D.P. editor. Populations of small mammals under natural conditions. Special Publication. Pymatuuning Laboratory of Ecology, University of Pittsburgh.

Frey, J.K. & Malaney, J.L. (2009) Decline of the Meadow Jumping Mouse (*Zapus hudsonius luteus*) in two Mountain Ranges in New Mexico. *The Southwestern Naturalist*, **54**, 31-44.

Galewski, T.M., Sanchez, S., Paradis, P., & Douzery, E.J. (2006) A Molecular Perspective on Mammalian Evolution from the Gene Encoding Interphotoreceptor Retenoid Binding Protein, with Convincing Evidence or Bat Monophyly. *Molecular Phylogenetics and Evolution,* **1**, 148-160.

Getz, L.L. (1985) Habitats. *Biology of New World Microtus.* (ed. by R.H. Tamarin) pp. 286-309. American Society of Mammalogist. Shippensburg, PA.

Gilpin, M.E., & Soule, M.E. (1986) Minimum viable populations: Processes of species extinction. Pages 19-34 in M.E. Soule, editor. Conservation Biology. Sinauer Associates, Sunderland, Massachusetts.

Golenishchev, F.N., Malikov, V.G., Nazari, F., Vaziri, A.S., Sablina, O.V. & Polyakov, A.V. (2002) New species of vole of "guentheri" group (Rodentia, Arvicolinae, *Microtus*) from Iran. *Russian Journal of Theriology,* **1**, 117-123.

Gromov, I.M. & Polyakov, I.Y. (1992) *Voles (Microtinae).* Onyx Publishers, New Delhi, India.

Guindon S & Gascuel O (2003). A simple, fast and accurate method to estimate large phylo- genies by maximum-likelihood". *Systematic Biology,* **52**, 696-704.

Grummer, J., Bryson, R.W., & Reeder, T.W. (2014) Species Delimitation Using Bayes Factors Simulations and Application to the *Sceloporus scalaris* Species Group (Squamata: Phrynosomatidae). *Systematic Biology*, **63**, 119-133.

Hall, E. R. (1981) *The mammals of North America.* John Wiley & Sons. New York, NY, USA.

Hellborg, L. (2004) Evolutionary studies of the mammalian Y chromosome. PH.D dissertation, Uppsala University.

Helend, J., & Drummond, A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.

Hoffmann, R.S., & Koeppl, J.W. (1985) Zoogeography. *Biology of New World Microtus.* (ed. by R.H. Tamarin) pp. 85-115. American Society of Mammalogist. Shippensburg, PA.

Hope, A.G., Waltari, E., Dokuchaev, N.E., Abramov, S., Dupal, T., Tsvetkova, A., Henttonen, H., MacDonald, S.O. & Cook, J.A. (2010) High-latitude diversification within Eurasian least shrews and Alaska tiny shrews (Soricidae). *Journal of Mammalogy,* **91**, 1041-1057.

Hope, A.G., Panter, N., Cook, J.A., Talbot, S.L., & Nagoren, D.W. (2014) Multilocus phylogeography and systematic revision of North American water shrews (genus: *Sorex*). *Journal of Mammalogy*. **95**, 722-738.

Hope, A.G., Waltari, E., Malaney, J.L., Payer, D.C., Cook, J.A. & Talbot, S.L. (2015) Arctic biodiversity: increasing richness accompanies shrinking refugia for a cold-associated tundra fauna. *Ecosphere*, **6**, 1-67.

Hope, A.G., Malaney, J.L., Bell, K.C., Salazar-Miralles, F., Chavez, A.S., Barber, B.R., & Cook J.A. (2016) Revision of widespread red squirrels (genus: *Tamiasciurus*) highlights the complexity of speciation within North American forests. *Molecular Phylogenetics and Evolution,* **100**, 170-182.

Hotaling, A.S., Percival, H.F., Kitchens, W.M., & Kasbohm, J.W. (2010) The Persistence of Endangered Florida Salt Marsh Voles in Salt Marshes of the Central Florida Gulf Coast. *Southeastern Naturalist,* **9**, 795-802.

Hubbard, J.P., Thaeler Jr., C.S., & Schmitt, C.G. (1983) Notes on voles (*Microtus*, Family Cricetidae) in New Mexico. *Occasional Papers of the Delaware Museum of Natural History,* **28**, 1-9.

Ikeda, H., Carlsen, T., Fujii, N., Brochmann, C., & Setoguchi, H. (2012) Pleistocene climatic oscillations and the speciation history of an alpine endemic and a widespread arctic-alpine plant. *New Phytologist*, **194**, 583-594.

Irwin, D. M., Kocher, T.D., & Wilson, A.C. (1991). Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution,* **32**, 128-144.

Jaarola, M. & Searle, J.E. (2002) Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, **11**, 2613-2621.

Jaarola M., Martínková N., Gündüz İ., Brunhoff C., Zima J., Nadachowski A., Amori G., Bulatova N.S., Chondropoulos B., Fraguedakis-Tsolis S., González-Esteban J., López-Fuster M.J., Kandaurov A.S., Kefelioğlu H., da Luz Mathias M., Villate I. & Searle J.B. (2004) Molecular phylogeny of the speciose vole genus *Microtus* (Arvicolinae, Rodentia) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution,* **33**, 647–663.

Jones, A.K., Liphardt, S.W., Perry, T.L. & Cook, J.A. In prep. Mammals of the greater

Gila region. Occasional Papers of the Museum of Southwestern Biology.

Kass, R.E., & Raftery, A.E. (1995) Bayes factors. *Journal of American Statistical Association*. **90**, 773-795.

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

Kohli, B.A, Fedorov, V.B., Waltari, E. & Cook, J.A. (2014) Phylogeography of a Holarctic rodent (*Myodes rutilus*): testing high-latitude biogeographical hypotheses and dynamics of range shifts, *Journal of Biogeography,* **42**, 377-389.

Koprowski, J.L., Ramos, N., Pasch, B.S., & Zugmeyer, C.A. (2006) Observations on the ecology of the endemic Mearns's squirrel (*Tamiasciurus mearnsi*). Southwest Naturalist, **51**, 426–430.

Krystufek, B., & Keflioglu, H. (2001) Redescription and species limits of *Microtus irani* Thomas 1921, and description of a new social vole from Turkey (Mammalia: Arvicolinae). Bonner Zoologische Beitrage, **50**, 1-14.

Krystufek, B., Buzan, E.V., Vohralik, V., Zareie, R., & Ozkan, B. (2009) Mitochondrial cytochrome *b* sequences yield new insight into the speciation of social voles in south-west Asia. *Biological Journal of the Linnean Society,* **98**, 121-128.

Krystufek, B., Zorenko, T., & Buzan, E.V. (2012) New insights into the taxonomy and phylogeny of social voles inferred from mitochondrial cytochrome *b* sequences. *Mammalian Biology,* **77**, 178-182.

Kumar, S., Stecher, G., & Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution.

Kurten, B., & Anderson, E. (1980) *Pleistocene mammals of North America*. Columbia University Press, New York, NY.

Laerm, J., Avise, J.C., Patton, J.C., & Lansman, R.A. (1982) Genetic determination of the status of an endangered species of pocket gopher in Georgia. *Journal of Wildlife Management*, **46**, 513-518.

Lienert, J., Fischer, M., & Diemer, M. (2002) Local extinctions of the wetland specialist *Swertia perennis* L. (Gentianaceae) in Switzerland: a revisitation study based on herbarium records. *Biological Conservation,* **103**, 65-76.

List, R., Pergams, O.R.W., Pachecho, J., Cruzado, J., & Ceballos, G. (2010) Genetic

divergence of *Microtus pennsylvanicus chihuahuensis* and conservation implications of marginal population extinctions. *Journal of Mammalogy*, **91**, 1093-1101

Lowery, G.H., Jr. (1974) *The mammals of Louisiana and its adjacent waters*. Louisiana State University Press, Baton Rouge, LA.

Lowry, M.L. (1998) Systematics of Northeastern Meadow Vole (*Microtus pennsylvanicus*) Subspecies, with Emphasis on the Island Endemic (*M. p. shattucki*, Howe 1901) in Penobscot Bay, Maine. Electronic Theses and Dissertations. Paper 648.

Lyons, L.A., Laughlin, T.A., Copeland, N.G., Jenkins, N.A., Womak, J.E., & O'Brien, S.J. (1997) Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nature Genetics*, **15**, 47-56.

MacDonald, S. O., Cook, J. A., Kirkland, G., Jr., & Yensen, E. (1998) *Microtus pennsylvanicus. North American rodents: Status survey and conservation action plan*.(eds. D.J. Hafner, E. Yensen, & G.L. Kirkland Jr.) pp. 99-101. Gland, Switzerland and Cambridge, UK.

Mace, G.M. (2004) The role of taxonomy in species conservation. *Philosophical Transactions of the Royal Society B*, **359**, 711-719.

Malaney, J.L. & Cook, J.A. (2013) Using biogeographical history to inform conservation: the case of Preble's meadow jumping mouse. *Molecular Ecology*, **22**, 6000-6017.

Markova E.A., V.P. Starikov, L.E. Yalkovskaya, S.V. Zykov, A.V. Morozkina, and P.A. Sibiryakov. 2014. Molecular and Cytogenetic Evidence for the Occurrence of the East European Vole *Microtus rossiaemeriionalis* (Arvicolinae, Rodentia) in the North of West Siberia. *Doklady Biological Sciences,* 455, 129-131.

Martin, R.A. (1968) Late Pleistocene distribution of *Microtus pennsylvanicus*. *Journal of Mammalogy*, **49**, 265-271.

Martinkova, N., Zima, J., Jaarola, M., Macholan, M., & Spitzenberger, F. (2007) The origin and phylogenetic relationships of *Microtus bavaricus* based on karyotype and mitochondrial DNA sequences. *Folia Zoologica*, **56**, 39-49.

Martinkova, N., & Moravec, J. (2012) Multilocus phylogeny of arvicoline voles (Arvicolini, Rodentia) shows small tree terrace size. *Folia Zoology,* **61**, 254-267.

Mayr, E. (1942) *Systematics and the origin of species*. New York: Columbia University Press.

McCleery, R.A., & Zweig, C.L. (2016) Leveraging limited information to understand

ecological relationships of endangered Florida salt marsh vole. *Journal of Mammalogy*. DOI:10.1093/jmammal/gyw084.

McCormack, J.E., Heled, J., Delaney, K.S., Peterson, A.T., & Knowles, L.L. (2011) Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. *Evolution,* **65**, 184-202.

Meyer, M.N., Golenishchev, F.N., Radjably, S.I. & Sablina, O.V. (1996) Voles (subgenus *Microtus* Schrank) of Russia and adjacent territories. *Proceedings of the Zoological Institute, Russian Academy of Sciences,* **232**, 1-330.

Montuire, S., Markova, E., & Aniskin, V. (2013) Exploring phylogeography and species limits in the Altai vole (Rodentia: Cricetidae). *Biological Journal of the Linnean Society*, **108**, 434-452.

Moyer, C.A., Adler, G.H., & Tamarin, R.H. (1988) Systematics of New England *Microtus*, with emphasis on *Microtus breweri*. *Journal of Mammalogy,* **69**, 782-794.

Park, R.A., Trehan, M.S., Mausel, P.W., & Howe, R.C. (1988) The effects of sea level rise on U.S. coastal wetlands. *Holcomb Research Institute Rep*, **142**, 1-60.

Petrova, T.V., Zakharov, E.S., Samiya, R. & Abramson, N.I. (2014) Phylogeography of the narrow-headed vole *Lasiopodomys* (*Stenocranius*) *gregalis* (Cricetidae, Rodentia) inferred from mitochondrial cytochrome *b* sequences: an echo of Pleistocene prosperity. *Journal of Zoological Systematics Evolutionary Research*, **53**, 97-108.

Plante, Y., Boag, P.T., & White, B.N. (1987) Macrogeographic variation in mitochondrial DNA of meadow voles (*Microtus pennsylvanicus*). *Canadian Journal of Zoology,* **67**, 158-157.

Platt, A.R., Woodhall, R.W., & George, A.L., Jr. (2007) Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. *BioTechniques,* **43**, 58-62.

Pruitt, W.O., Jr. (1968) Synchronous biomass fluctuations of some northern mammals. *Mammalia,* **32**, 172-191.

Puckett, E.E., Etter, P.D., Johnson, E.A. & Eggert, L.S. (2015) Phylogeographic Analyses of American Black Bears (*Ursus americanus*) Suggest Four Glacial Refugia and Complex Patterns of Postglacial Admixture. *Molecular Biology and Evolution*, **32**, 2338-2359.

De Queiroz, K. (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879-886.

Reding, D.M., Bronikowski, A.M., Johnson, W.E., & Clark, W.R. (2012) Pleistocene and ecological effects on continental-scale genetic differentiation in the bobcat (*Lynx rufus*). *Molecular Ecology,* **21**, 3078-3093.

Repenning, C.A., Fejfar, O., & Heinrich, W.D. (1990) Arvicolid rodent biochronology of the Northern. International symposium on evolution, phylogeny, and biostratigraphy of arvicolids (Rodentia, Mammalia). Pp. 385-418. Pfeil-Verlag, Prague, Czech Republic.

Robinson, M., Catzeflis, F., Briolay, J., & Mouchiroud, D. (1997) Molecular Phylogeny of Rodents, with Special Emphasis on Murids: Evidence for Nuclear Gene LCAT. *Molecular Phylogenetics and Evolution,* **42**, 423-434.

Ronquist F, & Huelsenbeck, R.J. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. **19**, 1574-1574.

Rovatsos, M.T. & Giagia-Athanasopoulou, E.B. (2012) Taxonomical status and phylogenetic relations between the "thomasi" and "atticus" chromosomal races of the underground vole *Microtus thomasi* (Rodentia, Arvicolinae). *Mammalian Biology,* **77**, 6-12.

Sawyer, Y. (2014) Living on the edge: exploring the role of coastal refugia and island biology in the Alexander Archipelago of Alaska. Ph.D. dissertation, University of New Mexico. Albuquerque, New Mexico.

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

Snell, R.R. & Cunnison, K.M. (1983) Relation of geographic variation in the skull of *Microtus pennsylvanicus* to climate. *Canadian Journal of Zoology,* **61**, 1232-1241.

Snyder, D.J. & Peterson, M.S. (1999) Life history of a peripheral population of blue-spotted sunfish *Enneacanthus gloriosus* (Holbrook), with comments on geographic variation. *American Midland Naturalist,* **141**, 345-357.

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

Stephens, M., Smith, N.J., & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978:989.

Stephens, M. & Donnelly, P. (2003) A comparison of Bayesian methods for haplotype reconstruction. *American Journal of Human Genetics*, **73**, 1162-1169.

Swenson, N.G., & Howard, D.J. (2005) Clustering of Contact Zone, Hybrid Zones, and Phylogeographic Breaks in North America. *American Society of Naturalist*, **166**, 581-591.

Tamarin, R.H. (1985) *Biology of New World Microtus*. American Society of Mammalogy, Shippensburg, PA.

Tougard, C., Brunet-Lecomte, R., Fabre, M. & Montuire, S. (2008) Evolutionary history of tow allopatric *Terricola* species (Arvicolinae, Rodentia) from molecular, morphological, and palaeontological data, *Biological Journal of the Linnean Society,* **93**, 309-323.

Tougard, C., Montuire, S., Volobouev, V., Markova, E., Contet, J., Aniskin, V. & Quere, J. Exploring phylogeography and species limits in the Altai vole (Rodentia: Cricetidae). *Biological Journal of the Linnean Society,* **108**, 434-452.

US Fish and Wildlife Service (1991) Endangered and threatened wildlife and plants; endangered status for the Florida Salt Marsh Vole. *Federal Register*, **56**, 1457-1459.

Weksler, M., Lanier, H.C. & Olson, L.E. (2010) Eastern Beringian biogeography: historical and spatial genetic structure of singing voles in Alaska. *Journal of Biogeography,* **37**, 414-1431.

Wilson D.E. & Reeder D.M. (2005) *Mammal species of the world*. John Hopkins University Press, Baltimore, MD.

Woods, C.A., Post, W., & Kilpatrick, C.W. (1982) *Microtus pennsylvanicus* (Rodentia: Muridae) in Florida: A Pleistocene relict in a coastal saltmarsh. *Bulletin of the Florida State Museum Biological Science,* **28**, 25-52

Woods, C.A. 1992. Endangered: Florida Saltmarsh Vole (*Microtus pennsylvanicus dukecampbelli*). Pp. 131-139, *In* S.R. Humphrey (Ed.) Rare and Endangered Biota of Florida Vol. 1 Mammals. University Press of Florida. Gainesville, FL. 392 pp.

Xie, W.G., Lewis, P.O., Fan, Y., Kuo, L., & Chen, M.H. (2011) Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematics Biology*, **60**, 150-160.

Youngman, P.M. (1967) Insular populations of the meadow vole, *Microtus pennsylvanicus*, from northeastern North America, with descriptions of two new subspecies. *Journal of Mammalogy,* **48**, 579-588.

Zink, R.M. & Barrowclough, G.F. (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107-2121.

Figures and Tables

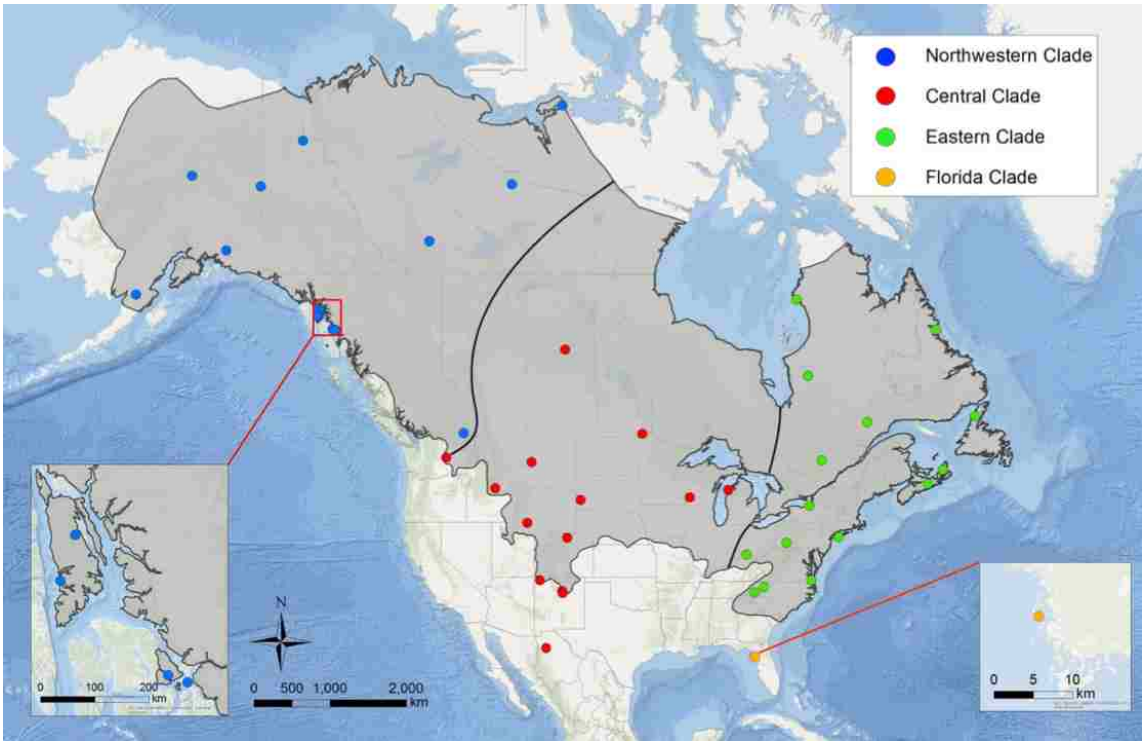


Figure 1: Distribution of *M. pennsylvanicus* is shown in gray (modified from IUCN) with dots representing sampling localities. Lines between major clades are hypothesized boundaries. Colors correspond to four distinct Cyt b clades with solid lines representing estimated range limits of each clade.

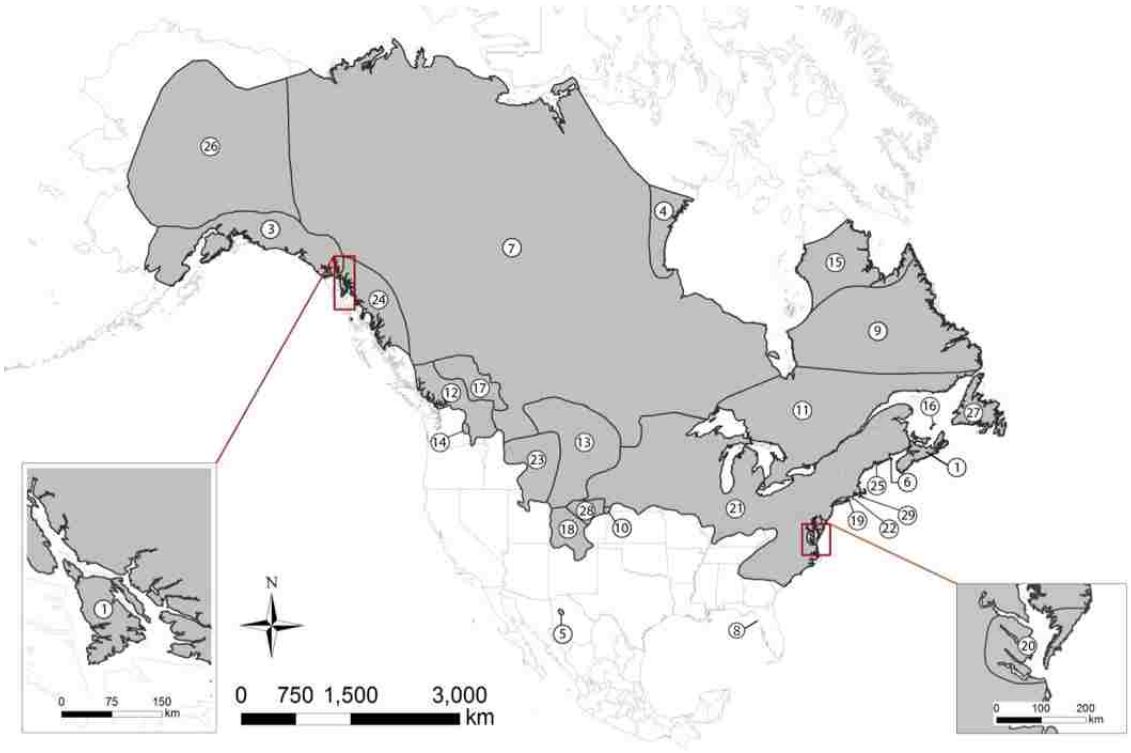


Figure 2: Distribution of *M. pennsylvanicus* is shown in gray (modified from Hoffmann and Koeppl 1985) and its insular allospecies *M. breweri*. Subspecies of *M. pennsylvanicus* are: 1, *M. p. acadicus*; 2, *M. p. admiraltiae*; 3, *M. p. alcorni*; 4, *M. p. alphorodemus*; 5, *M. p. chihuahuensis* (extinct); 6, *M. p. copelandi*; 7, *M. p. drummondii*; 8, *M. p. dukecampbelli*; 9, *M. p. enixus*; 10, *M. p. finitus*; 11, *M. p. fontigenus*; 12, *M. p. funebris*; 13, *M. p. insperatus*; 14, *M.p. kincaidi*; 15, *M. p. labradorius*; 16, *M. p. magdalensis*; 17, *M. p. microcephalis*; 18, *M. p. modestus*; 19, *M. nesophilus* (extinct); 20, *M. p. nigrans*; 21, *M. p. pennsylvanicus*; 22, *M. p. provectus*; 23, *M. p. pullatus*; 24, *M. p. rubidus*; 25, *M. p. shattucki*; 26, *M. p. tananaensis*; 27, *M. p. terraenovae*; 28, *M. p. uligocola*; 29, *M. breweri*

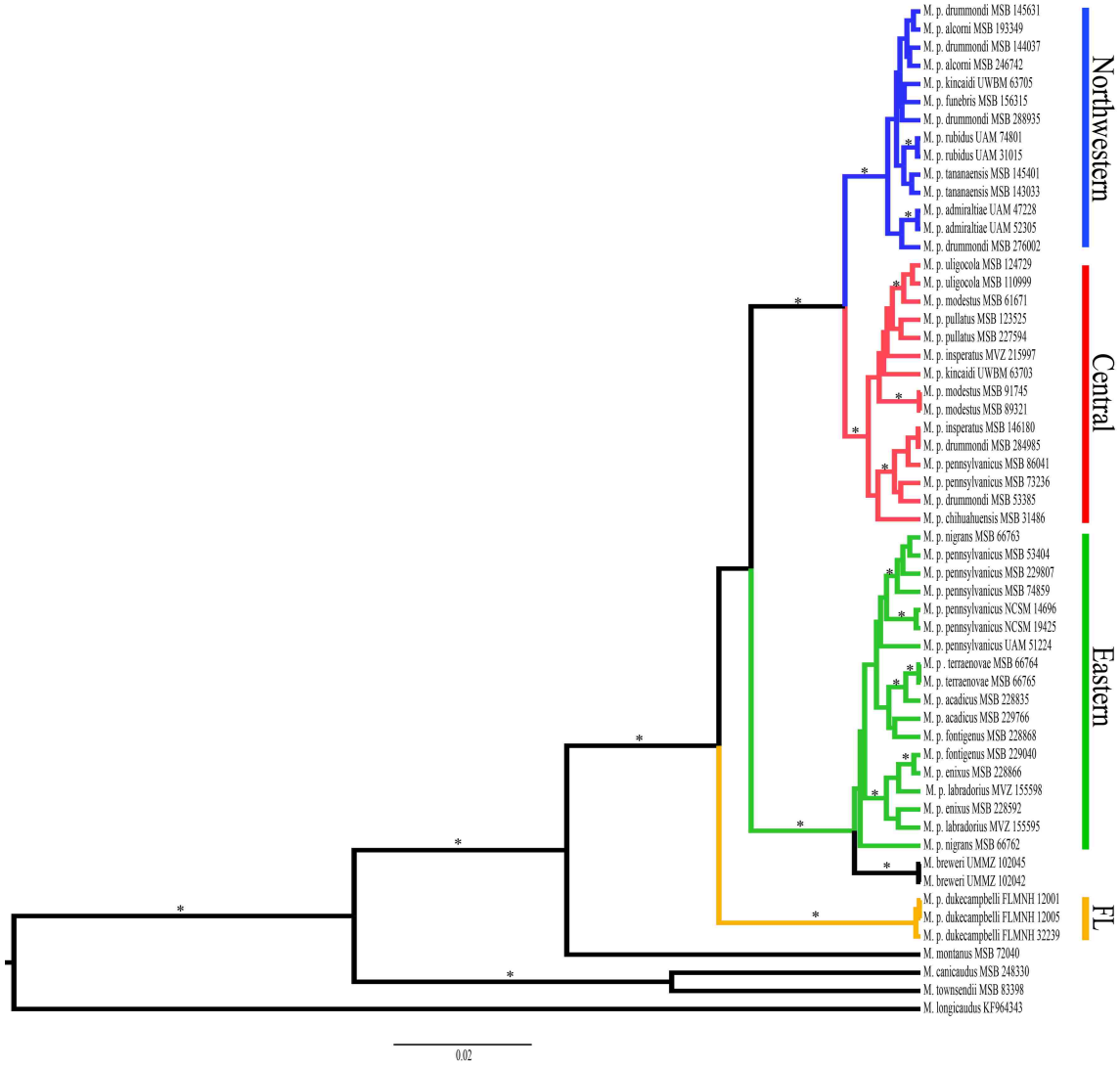


Figure 3: A Bayesian *Cyt b* gene tree for 48 individuals of *M. pennsylvanicus.* Colors correspond to *Cyt b* clade (Figure 1). Posterior probability of >0.95 are depicted by asterisk. Five species of *Microtus* were used as outgroups and are shown in black (*M. longicaudus, M. townsendii, M. canicaudus, M. breweri* and *M. montanus*).

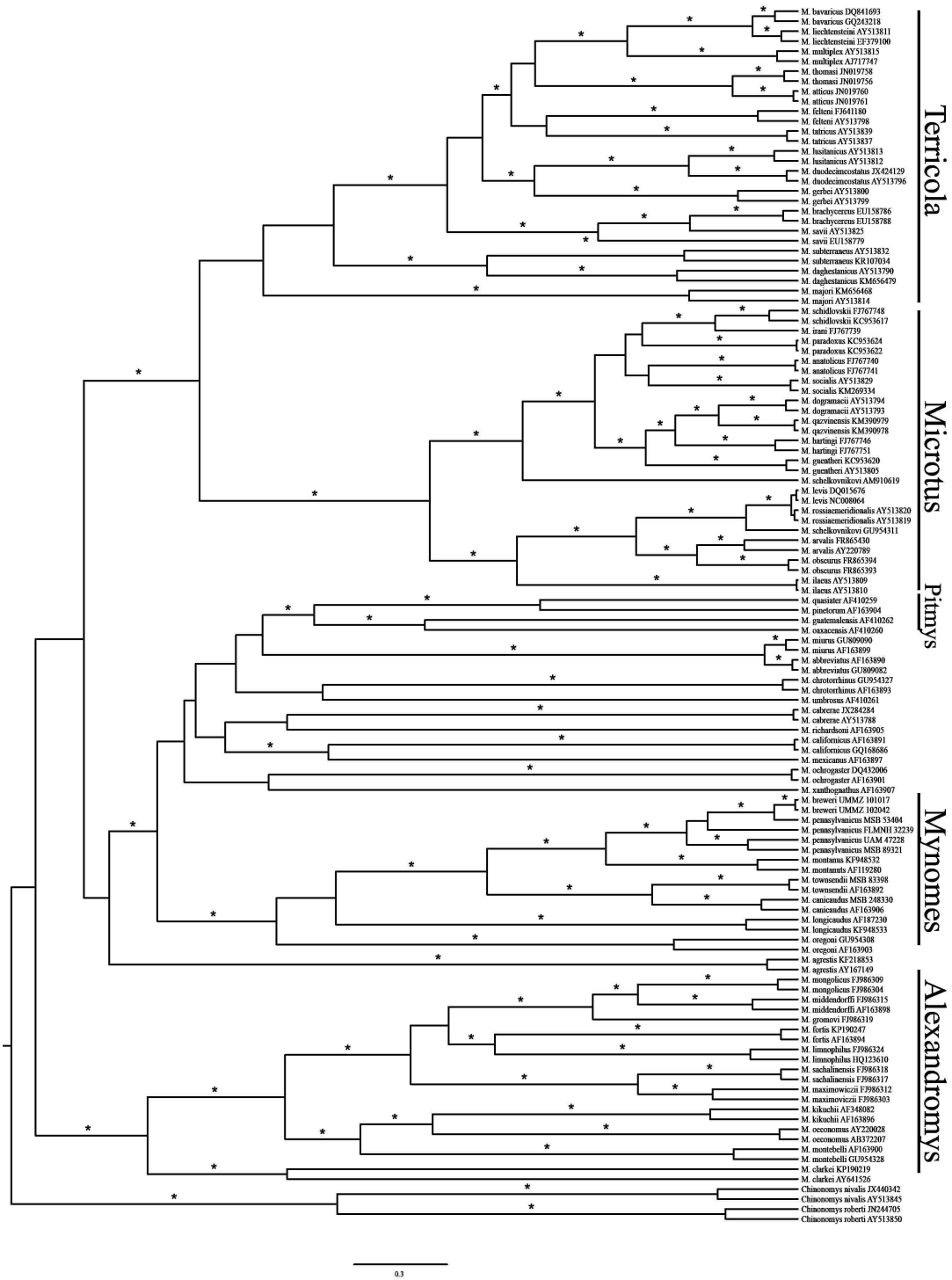


Figure 4: A Bayesian Cyt b gene tree for 63 species of *Microtus* with subgenera labeled vertically. Two species of *Chionomys* were used as outgroups and are shown at the bottom of the tree (*C. roberti* and *C. nivalis*).

Table 1. Estimates for diversity for each major mtDNA clade in *M. pennsylvanicus*. Numbers below the diagonal are the mean divergence and standard error estimates are shown in above diagonal and are based on 1,000 bootstrap replicates.

Clades	Northwestern	Central	Eastern	Florida	*M. breweri*
Northwestern	-	0.003	0.006	0.006	0.008
Central	0.018	-	0.005	0.006	0.008
Eastern	0.043	0.041	-	0.005	0.004
Florida	0.045	0.046	0.041	-	0.008
M. breweri	0.047	0.044	0.018	0.049	-

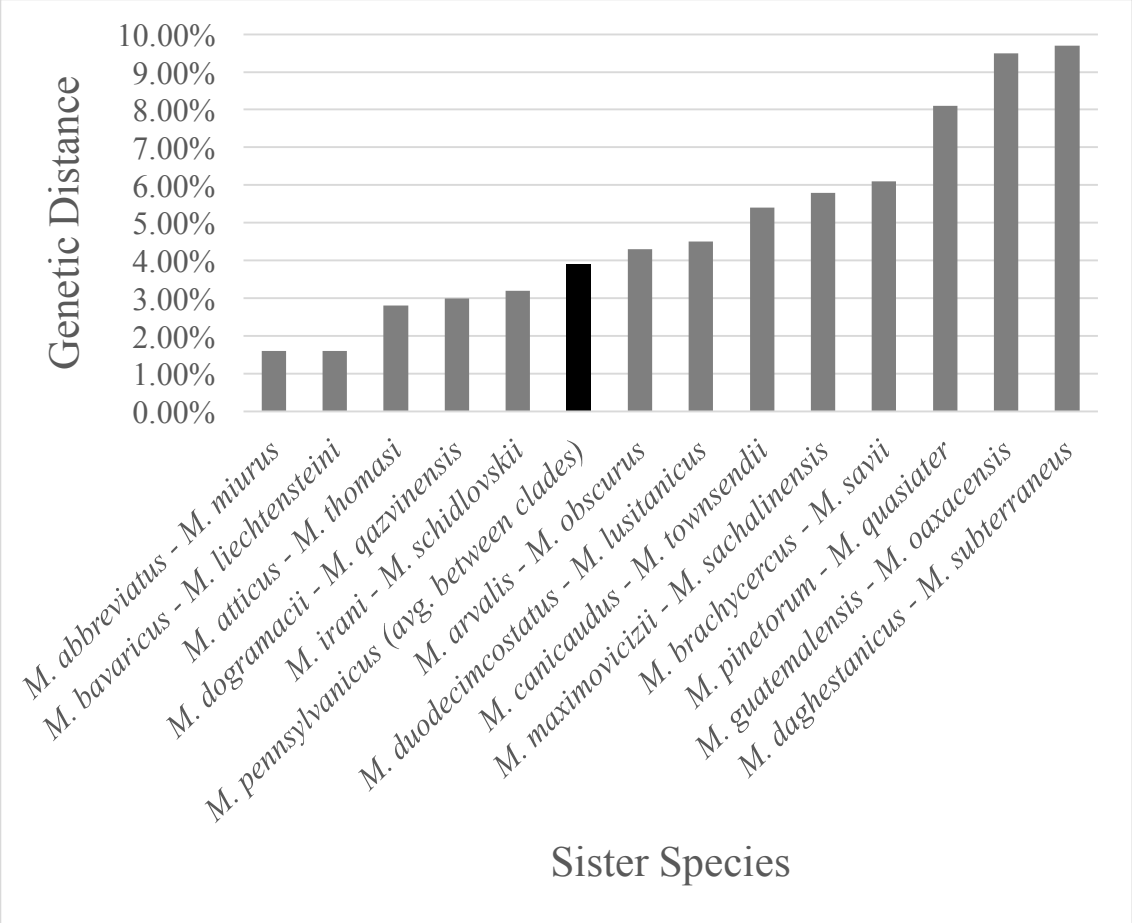
Table 2. Marginal likelihood estimates and Bayes factor testing results (2*ln*Bf) for alternative species-tree hypotheses of relationships among *Microtus pennsylvanicus* lineages. The marginal likelihood estimate with the best score is indicated by N/A and support for the best hypothesis over alternative hypotheses. Bayes factors are evaluated based on the recommendations of Kass and Raftery (1995), thus support for 2*ln*Bf of 0-2 = indicates "not worth more than mention", 2*lf*BF = 2-6 means "positive" support, 2*ln*Bf = 6-10 means "strong" support and 2*ln*Bf > 10 means "decisive" support in distinguishing between species delimitation hypotheses.

Model	All 7 loci				Six nuclear loci			
	PS		SS		PS		SS	
	MLE	2*ln*Bf	MLE	2*ln*Bf	MLE	2*ln*Bf	MLE	2*ln*Bf
H_1	-11507.6	105.28	-11508.05	105.14	-7889.41	37.62	-7889.62	37.22
H_{2a}	-11468.4	26.8	-11468.91	26.86	**-7870.6**	**N/A**	**-7871.01**	**N/A**
H_{2b}	-11491.8	73.6	-11492.24	73.52	-7884.7	28.2	-7885.13	28.24
H_3	-11463	16	-11463.27	15.58	-7871.3	1.4	-7871.64	1.26
H_4	**-11455**	**N/A**	**-11455.48**	**N/A**	-7872.74	4.28	-7873.3	4.58

Table 3. *Microtus* species with corresponding subspecies based on the geographic extent of each species range.

Genus	species	subspecies
Microtus	*drummondii*	*admiraltiae*
M.	*d.*	*alcorni*
M.	*d.*	*alphorodemus*
M.	*d.*	*chihuahuensis*
M.	*d.*	*drummondii*
M.	*d.*	*finitus*
M.	*d.*	*funebris*
M.	*d.*	*insperatus*
M.	*d.*	*kincaidi*
M.	*d.*	*microcephalis*
M.	*d.*	*modestus*
M.	*d.*	*pullatus*
M.	*d.*	*rubidus*
M.	*d.*	*tananaensis*
M.	*d.*	*uligocola*
M.	*dukecampbelli*	
M.	*pennsylvanicus*	*acadicus*
M.	*p.*	copelandi
M.	*p.*	enixus
M.	*p.*	*fontigenus*
M.	*p.*	*labradorius*
M.	*p.*	*magdalensis*
M.	*p.*	*nesophilus*
M.	*p.*	*nigrans*
M.	*p.*	*pennsylvanicus*
M.	*p.*	*provectus*
M.	*p.*	*shattucki*
M.	*p.*	*terraenovae*
M.	*p.*	*breweri*

Appendices



Appendix A. Genetic distance between highly supported sister species in Figure 4. Analyses were carried out in MEGA7 using a K2P model with pairwise deletions. The average genetic distance between four monophyletic clades within *M. pennsylvanicus* is shown in black with all other sister species considered shown in gray.

Appendix B. *Microtus* samples used in Figure 3. Included are taxonomic designation for each individual, the current status of the subspecies, museum ID, Museum voucher is located, mitochondrial clade and specimen was included in species tree analyses signified by "•". Subspecies designations follow Hoffmann and Koeppl (1985).

Genus	Species	Subspecies	Museum ID	mtDNA Clade	nuDNA Clade
Microtus	*pennsylvanicus*	*acadicus*	MSB 228835	E	
			MSB 229766	E	•
M.	*p.*	*admiraltiae*	UAM 47228	NW	•
			UAM 52305	NW	•
M.	*p.*	*alcorni*	MSB 193349	NW	
			MSB 246742	NW	
M.	*p.*				
M.	*p.*	*chihuahuensis*	MSB 31486	C	
M.	*p.*	*copelandi*			
M.	*p.*	*drummondii*	MSB 53385	C	•
			MSB 284985	C	•
			MSB 144037	NW	•
			MSB 143631	NW	•
			MSB 276002	NW	•
			MSB 288935	NW	•
M.	*p.*	*dukecampelli*	FLMNH 32239	FL	•
			FLMNH 12001	FL	
			FLMNH 12005	FL	
M.	*p.*	*enixus*	MSB 228592	E	•
			MSB 228866	E	•
M.	*p.*	*finitus*			
M.	*p.*	*fontigenus*	MSB 228868	E	
			MSB 229040	E	
M.	*p.*	*funebris*	MSB 156315	NW	•
M.	*p.*	*insperatus*	MSB 146180	C	•
			MVZ 215997	C	
M.	*p.*	*kincaidi*	UWBM 63703	C	
			UWBM 63705	NW	
M.	*p.*	*labradorius*	MVZ 155595	E	
			MVZ 155598	E	•
M.	*p.*	*magdalensis*			
M.	*p.*	*microcephalus*			

M.	*p.*	*modestus*	MSB 61671	C	
			MSB 89321	C	•
			MSB 91745	C	
M.	*p.*	*nesophilus*			
M.	*p.*	*nigrans*	MSB 66762	E	
			MSB 66763	E	
M.	*p.*	*pennsylvanicus*	MSB 53404	E	•
			MSB 73236	C	•
			MSB 74859	E	•
			MSB 86041	C	•
			MSB 229807	E	•
			NCSM 14696	E	•
			NCSM 19425	E	•
			UAM 51224	E	•
M.	*p.*	*provectus*			
M.	*p.*	*pullatus*	MSB 123525	C	•
			MSB 227594	C	•
			UAM 31015	NW	•
			UAM 74801	NW	•
M.	*p.*	*shattucki*			
M.	*p.*	*tananaensis*	MSB 143033	NW	
			MSB 145401	NW	
M.	*p.*	*terraenovae*	MSB 66764	E	
			MSB 66765	E	
M.	*p.*	*uligocola*	MSB 110999	C	
			MSB 124729	C	
M.	*breweri*		UMMZ 101017	outgroup	
M.	*montanus*		MSB 72040	outgroup	•
M.	*canicaudus*		MSB 248330	outgroup	•
M.	*townsendii*		MSB 83398	outgroup	•
M.	*longicaudus*		MSB 157001	outgroup	•

Appendix C. GenBank accession numbers for specimens obtained for this study.

Species	GenBank Accession Number		
M. bavaricus	DQ841693	M. anatolicus	FJ767741
M bavaricus	GQ243218	M. socialis	AY513829
M. liechtensteini	AY513811	M. socialis	KM269334
M. liechtensteini	EF379100	M. dogramacii	AY513793
M . multiplex	AJ717747	M. dogramacii	AY513794
M . multiplex	AY513815	M. qazinensis	KM390978
M. tatricus	AY513837	M. qazinensis	KM390979
M. tatricus	AY513839	M. hartingi	FJ767746
M. atticus	JN019760	M. hartingi	FJ767751
M. atticus	JN019756	M. guentheri	AY513805
M. thomasi	JN019756	M. guentheri	AY953620
M. thomasi	JN019758	M. schelkovnikovi	AM910619
M. felteni	AY513798	M. levis	DQ015676
M. felteni	FJ641180	M. levis	NC008064
M. duodecimcostatus	AY513797	M. rossiaemeridionalis	AY513819
M. duodecimcostatus	JX424129	M. rossiaemeridionalis	AY513820
M. lusitanicus	AY513812	M. schelkovnikovi	GU954311
M. lusitanicus	AY513813	M. arvalis	AY220789
M. gerbei	AY513799	M. arvalis	FR865430
M. gerbei	AY513800	M. obscurus	FR865393
M. brachycercus	EU158786	M. obscurus	FR865394
M. brachycercus	EU158788	M. ilaeus	AY513809
M savii	AY513825	M. ileaus	AY513810
M savii	EU158779	M. guatemalensis	AF410262
M. daghestanicus	AY513790	M. oaxaensis	AF410260
M. daghestanicus	KM656479	M. pinetorum	AF163904
M. subterraneus	AY513832	M. quasiater	AF410259
M. subterraneus	KR107034	M. chrotorrhinus	AF163893
M. majori	AY513814	M. chrotorrhinus	GU954327
M. majori	KM656468	M. umbrosus	AF410261
M. schidlovskii	FJ767748	M. abbreviatus	AF163890
M. schidlovskii	KC953617	M. abbreviatus	GU809082
M. irani	FJ767739	M. miurus	AF163899
M. paradoxus	KC953622	M. miurus	GU809090
M. paradoxus	KC953624	M. ochrogaster	AF163901
M. anatolicus	FJ767740	M. ochrogaster	DQ432006

M. xanthognathus	AF163907
M. cabrerae	AY513788
M. cabrerae	JX284284
M. richardsoni	AF163905
M. montanus	KF948532
M. montanus	AF119280
M. canicaudus	AF163892
M. townsensii	AF163906
M. longicaudus	AF187230
M. longicaudus	KF948533
M. oregoni	AF163903
M. oregoni	GU954308
M. californicus	AF163891
M. californicus	GQ168686
M. mexicanus	AF163897
M. agrestis	AY167149
M. agrestis	KF218853
M. middendorffi	AF163898
M. middendorffi	FJ986315
M. mongolicus	FJ986304
M. mongolicus	FJ986309
M. gromovi	FJ986319
M. fortis	AF163894
M. fortis	KP190247
M. limnophilus	FJ986324
M. limnophilus	HQ123t610
M. maximoviczii	FJ986303
M. maximoviczii	FJ986312
M. sachalinensis	FJ986317
M. sachalinensis	FJ986318
M. kikuchii	AF163896
M. kikuchii	AF348082
M. oeconomus	AB372207
M. oeconomus	AY220028
M. montebelli	AF163900
M. montebelli	GU954328
M. clarkei	AY641526
M. clarkei	KP190219
Chinonomys nivalis	AY513845
Chinonomys nivalis	JX440342
Chinonomys robeti	AY513850
Chinonomys robeti	JN244705

Chapter 3

Multilocus phylogeography of *Microtus pennsylvanicus* reveals deep structure across North America

Abstract

The cyclical glacial events of the Pleistocene significantly influenced the geographic distribution and genetic structure of organisms today. A better understanding of how organisms responded to these historical biogeographic events aids in forecasting how they may respond to projected climatic shifts. We utilized comprehensive range-wide sampling and multilocus phylogeographic analyses (1 mtDNA and 6 nuDNA genes) combined with ecological niche modeling to further our understanding of how a wide spread North American mammal, the meadow vole (*Microtus pennsylvanicus*), responded to changing climates during the Quaternary. Additionally, we used population demographic analyses to characterize how independent lineages of meadow vole responded to historical biogeographic events. Our findings suggest that the dynamic evolutionary history of *M. pennsylvanicus* is illustrated by long term persistence in multiple distinct refugia with subsequent expansion to newly available habitat following the retreat of ice sheets. Multiple geographic regions of glacial-interglacial persistence were identified south of the Laurentide and Cordilleran ice sheets; a western refugium in the proximity of the southern Rockies and two eastern refugia, one east of the Appalachian Mountains and another small relict in coastal Florida. The refugium for the Northwestern clade remains unclear as to whether populations persisted in the coastal refugium of Southeast Alaska, Beringia or both. Finally, we identified signals of long

term persistence in a cryptic refugium in southeast Canada that will be clarified with more extensive sampling of the region. The effects of glacial-interglacial cycles preceding the LGM are elucidated in this phylogeographic study, as they played a key role in the diversification and geographic distribution of *M. pennsylvanicus.*

Donavan J. Jackson [a], Jason L. Malaney[b*] and Joseph A. Cook[a*]

[a]Department of Biology and Museum of Southwestern Biology. University of New Mexico, Albuquerque NM 87131, USA, [b]Department of Biology, Austin Peay State University, Clarksville, TN 37044, USA

[*]Co-authors for submission of a version of this chapter to a peer-reviewed journal

Introduction

The Pleistocene was a dynamic epoch filled with climatic and environmental fluctuations that played an integral role in the contemporary distributions and genetic diversity of plants and animals (Hofreiter & Stewart 2009). Repeated climatic oscillations impacted the geographic ranges of organisms during cooling and warming phases by forcing species to shift their distributions depending on their ability to adapt or move through the landscape (Hewitt 2004; DeChaine 2008). From multiple studies of co-distributed species in boreal North America, common phylogeographic patterns have emerged that are found across a number of taxa from plants to animals (Brunsfeld *et al.* 2001; Fedorov *et al.* 2007; Soltis *et al.* 2008). For example, the ranges of many northern organisms were reduced to one or more of a series of shared glacial refugia during cold periods of the Pleistocene. Animals and plants alike moved generally northward in

warmer periods and southward as temperatures cooled, with glacial-interglacial refugial or isolation phases often leading many species to diversify during the late Quaternary (Hewitt 1996). Using molecular data, we are gaining a better understanding of the demographic and distributional impact on organisms affected by glacial-interglacial cycling (Avise 2009). Furthermore, phylogeographic approaches allow us to identify geographic barriers that prevented gene flow and common colonization corridors that facilitated recolonization of organisms from glacial refugia following the retreat of ice sheets. These key historical attributes of contemporary northern populations set the stage for interpreting the underlying spatial and demographic processes that produced patterns of genetic diversity across space and time (Avise 2000). Finally, a clearer understanding of how organisms responded to climate change in the past provides a foundation for better management of species in light of contemporary climate change.

During the cold phases of the Pleistocene, the Laurentide and Cordilleran ice sheets covered much of the northern latitudes of North America and sea levels decreased up to 120 m, resulting in the exposure of considerable terrestrial surface area and providing new connections between land masses (Hofreiter & Stewart 2009). Terrestrial organisms responded to climate cycles individually; while some persisted in unglaciated regions, others tracked habitat conditions (Riddle 1996). Vertebrate populations located in the montane regions of western North America often experienced fragmentation that coincided with contraction and expansion of habitats along elevational gradients during climate cycles. Populations isolated for extended periods began to diverge from other conspecific populations, often resulting in the formation of distinct geographic lineages. In contrast, contemporary populations that have recently colonized previously glaciated

regions tend to exhibit low genetic variation; a common genetic signature of many species across vast areas of high latitudes due to recent and rapid expansion (Lessa *et al*. 2003; Avise 2000). Sometimes as distinct lineages shifted distributions during the warmer periods (*i.e.,* interglacial), they may have experienced contact with other independently evolving lineages that were expanding from other refugia. Multiple contact zones situated at the same location across multiple species due to physiographic features result in suture zones (Swenson & Howard 2005) and point to the possibility that common processes have produced similar evolutionary outcomes. To better understand how organisms responded to Pleistocene climate change, independent perspectives can be integrated from molecular and niche modeling approaches (Waltari *et al*. 2007). With contemporary warming climate and expanding human populations, an understanding of how species responded to changing environments in the past will provide a foundation for better managing and conserving biodiversity in the future (Parmesan 2006).

Phylogeographic studies of North American species have detected signatures of expansion, contraction and hybridization (Runck *et al.* 2009; Garroway *et al*. 2010), presumed glacial refugia that populations persisted in during cold periods (Reding *et al.* 2012, Hope *et al.* 2014), distinct geographically dispersed lineages (Demboski & Cook 2001; Burbrink *et al.* 2007), and cryptic, putatively new species (Conroy & Neuwald 2008; Crawford *et al.* 2011). Previous studies highlighted that glacial and interglacial refugia played an important role in the contemporary genetic structure and distribution of species today (Lessa *et al.* 2003, Shafer *et al.* 2010). For example, the process of hybridization can play an important role in speciation by increasing genetic diversity through admixture (Soltis 2003). Phylogeography also provided a basis for management

decisions by elucidating antiquated taxonomic nomenclature (Malaney & Cook 2013) and through predictive methods that allow us to target species most at risk due to climate change (Hope *et al.* 2013). Furthermore, we can now assess ecological divergence between taxa through niche modeling (e.g., Conroy & Neuwald 2008). By integrating molecular data and ecological data, we can investigate the mechanisms that may lead to incipient speciation through the complimentary perspectives that emerge from incorporating multiple datasets.

The meadow vole (*Microtus pennsylvanicus*) represents an excellent organism to study how small mammals respond to changing environments spatially and temporally. Associated with mesic meadows and marshes, this widespread species presumably recolonized northern North America during the Holocene. The oldest fossil of meadow voles is from Kanopolis, Kansas and dates to the late Irvingtonian (1.8-0.24 mya). Additional Late Pleistocene fossils are found in Texas, southern New Mexico, Arkansas (these regions no longer occupied by *M. pennsylvanicus*) and in Florida (Martin 1968; Smartt 1977) making it the most widely distributed microtine during the Pleistocene (Kurten & Anderson 1980). The current transcontinental distribution ranges from 28-68° N latitude. Found along the Arctic Ocean at its northernmost limit, the meadow vole is commonly encountered throughout Alaska and Canada, and extending south along the Rocky Mountains to northern New Mexico and south along the Appalachian Mountains to Georgia (Hoffman & Koeppl 1981; Figure 1). In addition to this expansive distribution, several peripheral populations may warrant special conservation consideration including multiple Pacific island populations in Southeast Alaska. Whether these western peripheral-isolates are the product of sustained occupation through multiple

glacial-interglacial cycles or recent colonization is unclear. Second, there are at least three southern regions with peripheral isolates including a population occupying western Florida that is listed as federally Endangered (Federal Register 1991); a single population in northern Mexico that recently went extinct (List *et al*. 2010); and several populations in southern and central New Mexico that are likely extinct as they have not been detected in nearly a century (Ligon 1915; Bailey 1932; Hubbard *et al.*1983; this study). Finally, there exist a series of populations along the Atlantic Coast of eastern Canada and New England. One population is restricted to Muskeget Island, near the coast of Nantucket, Massachusetts, but two other nearby populations have gone extinct. All three of these are closely related to *M. pennsylvanicus*. One is considered a separate species (the beach vole, *Microtus breweri*) due to island effects and craniodental differences, while the other two populations (the Gull Island vole, *M. p. nesophilus*) are a subspecies. Fossil evidence suggests *M. pennsylvanicus* was once more widespread than the current range extents indicate, especially in the south. Consequently, it remains unclear if these multiple peripheral isolates (west, south, and east) have 1) persisted through multiple glacial-interglacial cycles and so are long-term relicts, 2) relicts that have been isolated since the Last Glacial Maximum (LGM), or 3) the product of very recent long-distance colonization.

Previous phylogeographic studies of organisms with North American distributions comparable to *M. pennsylvanicus* identified similar patterns of geographic structure that included multiple genetically distinct lineages (Dragoo *et al*. 2006; Burbrink *et al.* 2007; Puckett *et al.* 2015). In addition to regional variation, these widespread species often show similar signatures of northern expansion during the Holocene, suggesting glacial

events played a key role in structuring genetic diversity across these widespread species (Hewitt 1996, 2000). As a significant portion of the higher latitudes of the current range of *M. pennsylvanicus* was glaciated during the Quaternary, those regions should have predictable DNA signatures indicative of post-Pleistocene recolonization events following recession of the ice sheets (*e.g.*, Arbogast 1990; Conroy & Cook 2000). In contrast, southern populations that remained in unglaciated regions generally show signatures of population stability due to their longer period of persistence (Lessa *et al.* 2003). Phylogeographic investigations of several species of *Microtus* have provided key insight into how the dynamic climate of the Late Pleistocene influenced species distributions (Martinkova *et al.* 2007; Tougard *et al.* 2008; Haring *et al.* 2011).

In this phylogeographic study, we assess signatures of evolutionary and ecological processes that likely impacted the biogeographic history and shaped geographic structure within a widespread North American species. Because the majority of the contemporary range of *M. pennsylvanicus* was repeatedly covered by ice sheets during the Pleistocene, we predict that northern populations should reflect significant demographic growth and range expansion. However, the source populations of major colonization events remain largely unresolved for this species. A limited fossil record of *M. pennsylvanicus* in Beringia during the Quaternary, suggests that this species failed to occupy the high-latitudes north of ice sheets throughout the Pleistocene. Instead, we hypothesize that *M. pennsylvanicus* occupied multiple refugia south of the ice sheets and possibly in the Alexander Archipelago during the LGM. We will examine molecular variation using a multilocus approach and range-wide sampling to better understand how the Late Pleistocene affected the contemporary distribution of the meadow vole and to

further characterize presumed glacial refugia, colonization routes following glacial retreat, and timing of divergence between genetically distinct lineages. By exploring the demographic histories of *M. pennsylvanicus* lineages, we can begin to elucidate how environmental change led to expansion and contraction of these arvicoline rodents through time. We hypothesize that following the retreat of the Laurentide and Cordilleran ice sheets, *M. pennsylvanicus* moved into newly available habitat in the North. We expect to see a genetic signal of rapid expansion in populations in high latitudes while populations south of the ice sheets will not show demographic expansion. Finally, we will utilize niche modeling to assess ecological divergence between geographically distinct lineages.

Materials and Methods

Sampling and Sequencing

We obtained 1 to 3 specimens (n = 148 total) from 63 localities that were spread throughout the geographic distribution of *M. pennsylvanicus* (Figure 1) representing 20 of 28 subspecies (Hoffman & Koeppl 1985). Tissues were obtained from seven museums and for all the mtDNA cytochrome b gene was sequenced. A subset of 33 individuals that represented major clades had 6 nuclear DNA loci (nuDNA) sequenced for a total of 5,025 base pairs (bp) of combined mtDNA and nuDNA sequence data. To increase geographic breadth, 22 partial mtDNA sequences from previous studies of *M. pennsylvanicus* were downloaded from GenBank. Our dataset included mitochondrial Cytochrome b (Cytb; 363-1140 bp), and nuclear sequences including Protein C-est-2 (ETS2; 875 bp), beta-fibrinogen (FGB; 600 bp), growth hormone receptor (GHR; 460 bp), interphotoreceptor

retinoid binding protein (IRBP; 625), lecithin cholesterol acyl transferase (LCAT; 465), and recombination-activating protein 1 (Rag1; 860 bp). For mtDNA we included 3 individuals from the allospecies *M. breweri* and 2 individuals from an additional four closely related species (*M. canicaudus, M. longicaudus, M. montanus,* and *M. townsendii*) to root the phylogeny (4 from GenBank and 4 newly sequenced). For nuDNA analyses we were unable to sequence nuclear loci for *M. breweri* so only outgroups were used in phylogenies.

Total genomic DNA was extracted from liver (frozen at -80 °C or ethanol preserved), muscle (ethanol preserved), or skin clips. All skin clips and ethanol preserved samples were washed in STE buffer overnight before extraction. We used a standard salt extraction or QIAamp DNA Mini Kit (Qiagen Inc., Valencia, California) and subsequently amplified mtDNA and nuDNA through polymerase chain reaction (PCR) and cycle sequencing. Primers pairs used for double stranded amplification of Cytb included MSB05/MSB14, and for degraded DNA the internal primer pairs were: MSB05/MVZ04, MSB11/MVZ26, L649/H885, and L820/MSB14. Primer pairs and corresponding annealing temperatures for nuclear loci are listed in Table 1. Reagents and PCR conditions can be found in supplemental information.

Sequences were edited and aligned using the MUSCLE algorithm (Edgar, 2004) in Geneious v.8.0.5 (Kearse et al., 2012) and visually confirmed. All sequences were blasted to GenBank to ensure that sequences corroborated species identification. To confirm that we had genuine mtDNA sequences (DeWoody *et al*., 1999), we translated all nucleotide bases to amino acids to confirm that no internal stop codons were present. Furthermore, transversion/transition changes and the frequency of the 1st, 2nd, and 3rd

codon positions were examined and no anomalies were found. We inferred nuclear alleles from heterozygotes using PHASE v2.1.1 (Stephens *et al.*, 2001). Each gene was run 4 times and the best goodness-of-fit to an approximate coalescent model was retained. Results were only retained with a >90% probability; genes with a lower probability were retained as N in sequence data to prevent bias. The accepted run resulted in two phased nuclear haplotypes (alleles) per individual and for subsequent analyses, one randomly chosen allele was utilized.

Gene Tree Reconstruction

To maximize geographic sampling within the Cytb dataset, we included all available partial and complete sequences. The best DNA model of substitution for each gene was determined in jModelTest v.2.1.7 (Darriba *et al*. 2012; Guindon & Gascuel 2003) under the Bayesian Information criterion. For phylogeny reconstruction, we used Bayesian methods and performed Markov Chain Monte Carlo searches in BEAST v1.8.2 (Drummond *et al.* 2012). For the Cytb gene tree construction two independent runs were computed for 20 million generations, sampling every 2000 generations with a burnin of 20%. The mutation rate for Cytb was set to 7.72×10^{-2} substitutions/site/Myr following Hope *et al*. (2014). Nuclear genes were computed at 15 million generations and were optimized to either a strict or lognormal relaxed clock. Log and tree files were combined using LogCombiner (software included with BEAST package) and stationarity of MCMC runs were assessed using Tracer ensuring Effective Sample Size (ESS) were above 200 (Rambaut & Drummond 2007). Final tree was visualized in FigTree v1.4.2 (Rambaut 2009).

Species trees

Files for species trees were formatted using BEAUti, a software package included in BEAST v.1.8.2 (Drummond & Rambaut 2007). We used the *BEAST algorithm to conduct a MCMC multilocus coalescent search for the species tree relationships (Heled & Drummond 2010). Gene trees constructed individually may have discordance due to distinctive gene histories; however, to resolve incongruence we used multiple independent loci to estimate a species tree. We included mitochondrial (n=1) and nuclear (n=6; Table 1) loci to provide independent perspectives of the evolutionary history of *M. pennsylvanicus*. For nuclear loci, we included a subset of individuals from each mitochondrial clade identified from preliminary mtDNA analysis. Additionally, we sampled individuals from multiple geographic localities within each *M. pennsylvanicus* clade to maximize geographic breadth. For each dataset all loci were unlinked across all parameters and priors for models of evolution were inferred from jModelTest (Darriba *et al.* 2012; Guindon & Gascuel 2003). Samples were assigned to clades determined by Cytb phylogeny. The mutation rate prior used for Cytb gene tree analyses was used for the species tree and all nuclear loci were estimated to be relative to prior. Proper ploidy was ensured for all nuclear and mitochondrial loci. Bayes Factor test revealed that there was no significant difference between clock models, indicating all loci were evolving at a relatively constant rate, therefore we used a strict clock for all loci. We used a Yule tree prior with piecewise linear and constant root was applied. Two independent runs were carried out, each was run for 250 million generations, sampling every 5,000 generations. Log files were assessed in Tracer (Rambaut & Drummond 2007) ensuring all ESS values were above 200.

Population Genetic Inference and Demographic Analyses

For all demographic analyses a subset of the Cytb dataset were utilized to maximize sampling and sequence length (n=148 and 1,111 bp, respectively). Individuals were assigned to groups for population analyses based on inclusion in highly supported Cytb clades. Genetic distance within and between all clades was computed in MEGA7 (Kumar *et al.* 2015) using the Kimura 2-parameter model, and standard errors were bootstrapped 1,000 replicates with pairwise deletions for missing data. Genetic diversity and demographic analyses were examined in DnaSP v5.10 (Librado & Rozas 2009); we calculated the sample size (N), number of segregating sites (S), number of haplotypes (*H*), haplotype diversity (*Hd*), mean number of nucleotide differences (k) and nucleotide diversity (π) for each clade. Additionally, each clade was divided between regions covered by ice during the LGM and regions that remained ice free to investigate signatures of population expansion. Analyses of molecular variance (AMOVA) were carried out in Arlequin v2.0 (Schneider *et al.* 2000) to estimate the levels of population structure in *M. pennsylvanicus* at various geographical scales.

We investigated historical population size changes using the Extended Bayesian Skyline algorithm that is implemented in BEAST for each widespread Cytb lineage (Heled and Drummond 2008). Subsequently, we tested for signals of population disequilibrium in three separate hierarchal datasets using population genetic methods: Fu's F_s (Fu 1997), R2 (Ramos-Onsins & Rozas 2002) and Tajima's *D* (Tajima 1989). Each were carried out independently in DnaSP using 10,000 coalescent simulations to test for significance. Those tests belong to separate classes of neutrality (Ramos-Onsins & Rozas 2002) and were applied to highly supported mtDNA clades and subsequently each clade split between glaciated and unglaciated regions during the Pleistocene.

Additionally, we tested each nuclear locus for signals of recent demographic expansion. A significantly negative Fu's F_s and Tajima's *D* along with a small positive R2 values are indicative of population growth in the past, consistent with a scenario of expansion. Fu's F_s uses haplotype data and rare alleles while R2 and Tajima's *D* use segregating sites to infer demographic histories based on nucleotide data. Lastly, we constructed mismatch distributions using a goodness-of-fit test to determine if the observed data are consistent with a model of recent expansion (Rogers & Harpending 1992). Analyses were run for 10,000 iterations on each major clade and plotted because different demographic histories will exhibit distinctive patterns using pairwise differences.

Ecological Niche Modeling

Bioclimatic envelope modeling

We used bioclimatic envelope modeling (BEM) (Araújo & Peterson 2012) to characterize occupied distributions and reconstruct paleodistributions of each lineage/species. We used these niche-based approaches as an independent perspective of ecological divergence among putative species using a two-step approach. First, we reconstructed contemporary distributions and then projected (transferred) the final model to ancestral climate conditions (*i.e.*, LGM). Then, we assessed spatial overlap of predicted distributions.

Occurrence and environmental data

BEMs are typically constructed using two forms of data: occurrence records and environmental layers. We downloaded records of *M. pennsylvanicus* from VertNet (accessed June 2016) and conducted a series of screening and filtering steps to reduce bias. First, we removed all samples with a georeferenced error >5km, those that failed to

provide an estimate of error, and those that were outside the known distribution of the species. Next, to reduce spatial autocorrelation, we used a custom aggregated reduction technique by retaining a single random sample within $10km^2$ radius and discarding other spatially redundant locations (Fourcade *et al.* 2014; Warren *et al.* 2014).

To assess lineage-based distributional overlap and test for recent niche divergence among evolutionarily independent but closely related groups we used the mtDNA haplotypes to quantify areas occupied by lineages. First, we added a minimum convex polygon in ArcGIS v10.2 to each linage-based set of points and applied a 2.5 decimal degree buffer (~250 km) and assigned all points within each polygon to the respective mtDNA lineage. We then reassigned mtDNA samples to putative species identified from the best model in tests of evolutionary hypotheses.

The geographic extent of species is an important consideration when using correlative models (Barve *et al.* 2011). We applied an extent based on the known geographic distribution of *M. pennsylvanicus*. Next, we generated a bias file to optimize background and occurrence point selection within the study extent. Bias files are frequently used in correlative modeling to avoid oversampling and can offset effects of geographic biases often associated with coordinate-based data, decreasing commission (false-positive) error rates (Anderson & Raza 2010, Barve *et al.* 2011, Merow *et al.* 2013). We used SDMToolBox v1.1c (Brown, 2014) to generate the bias files for each species, using the "Sample by Buffered Local Adaptive Convex-Hull" tool that limits background points (Thuiller *et al.* 2009; Barbet-Massin *et al.* 2012) and set the buffer distance to 75km and the alpha parameter to four.

We used 19 bioclimatic (temperature and precipitation) variables from the WorldClim database to quantify the ecological tolerances of each lineage of *M. pennsylvanicus*. These landscape-level data are often useful for assessing both current and paleodistributions (Waltari *et al*. 2007; Waltari & Guralnick 2009; Malaney & Cook 2013), but are also suitable for assessing Grinnellian-based niche divergence (McCormack *et al*. 2010).

Correlative modeling

We used MaxEnt version 3.3.3e (Phillips & Dudik 2008) due to superior performance over other correlative modeling approaches (Elith *et al*. 2006; Peterson *et al*. 2007) and ability to provide statistical comparisons among models (Elith *et al*. 2011; Merow *et al*. 2013). MaxEnt uses presence data in comparison with random background samples to estimate species distributions (Elith *et al*. 2011; Merow *et al*. 2013; Guillera-Arroita *et al*. 2015). MaxEnt is optimized with multiple default settings that require testing prior to modeling to enhance species-specific model performance (Anderson & Gonzalez 2011). Consequently, we conducted criterion-based model selection (Burnham & Anderson 2002; Warren & Seifert 2011) by assessing different combinations of the feature class types (FC) and regularization multipliers (RM) using ENMTools v.1.4.4 (Warren *et al*. 2010). We constructed single models for each species by applying alternate FCs (e.g., L - linear; LQ - linear and quadratic; H – hinge; LQH - linear, quadratic, and hinge; LQHPT – LQH, product, and threshold) and alternate RMs ranging from 0.01 to 5.0 at 0.5 intervals. We identified the optimum model settings using corrected Akaike information criterion (AICc). In sum, we tested 55 alternate models for each species (5 FCs x 11 RMs).

With the best-fit model parameters applied, we then conducted final modeling for each species using 20 replicates, log-scale outputs, and set the number of iterations to 5k but kept the default convergence threshold (= 0.00001). During replicate models, we applied the bias file for background sampling, retained 20% of presence localities as a training dataset, and used subsampling run type (jackknife approach) because of small sample sizes for some mtDNA lineages (Pearson *et al*. 2007; Shcheglovitova & Anderson 2013). We examined the standard deviations across all replicates for anomalous model behavior, which were minor.

MaxEnt produces a set of continuous surfaces and we used the mean of replicated models to represent relative suitability/likelihood (Richmond *et al*. 2010; Guillera-Arroita *et al*. 2015). To assess the difference between high suitability for each species, we constructed a single raster by calculating the difference between the continuous mean logistic outputs of comparative lineages (DiffAB = species A – species B). Further, we created a binary distribution of suitable and unsuitable areas using the equal training sensitivity and specificity logistic threshold for the spatial projection of final models (Pearson *et al*. 2007) and used the threshold-applied distributions to characterize paleodistributions and assess distributional shifts (see below).

Paleodistributions

We reconstructed paleodistributions of individual lineages to understand the history of geographic distributions, frequently yielding insights of ancestral areas as well as regions of ancestral admixture. Specifically, we projected predictions for the late-Pleistocene (*i.e*., LGM). Considering many North American mammals have shifted distributions (Lessa *et al*., 2003), we expanded the spatial extents through past

timeframes by 500 km to include areas nearby that might have been historically occupied by ancestral populations. Further, we validated paleodistribution reconstructions with fossil records obtained from FaunMap database (accessed June, 2016). We then assessed spatial overlap among clades using techniques analogous to contemporary overlap measurements.

Results

Mitochondrial DNA

The aligned Cytb dataset contained 160 sequences (144 new and 16 from GenBank) for *M. pennsylvanicus*. Partial sequences ranged from 379-1140 bp; however, only 5 individuals were shorter than 900 base pairs. The dataset was comprised of 176 variable sites, 108 parsimony informative sites and no stop codons were detected. Variation across codon positions is typical of mammalian mtDNA (Irwin *et al.* 1991), no insertions or deletions were identified and most variation was located in the third codon with the least variation at the second codon. The mean genetic distance within all *M. pennsylvanicus* sequences is 2.3%.

Bayesian analyses depict 4 well supported clades that are geographically structured (Figure 1); however, relationships between clades remains uncertain due to low branch support. The Northwestern clade (n=71) is located throughout Alaska and the northwestern provinces of Canada (Yukon, British Columbia, the Northwest Territories and Nunavut). The Central clade (n=46) occupies much of the region east of the Rockies and west of the Appalachian Mountains and found south into Mexico and north into Saskatchewan. The Eastern clade (n=39) inhabits regions along the eastern seaboard of

the Appalachian Mountains from northern Quebec to Georgia. Finally, the Florida clade (n=3) is extremely restricted and has only been identified in a limited number separate localities in the salt marshes along the Florida Gulf coast (Hoatling *et al.* 2010; McCleery & Zweig 2016; Figure 2).

Three of these clades are broadly distributed and largely parapatric. The Northwestern and Central clades narrowly come into contact in Washington and along the British Columbia and Alberta border (Figure 2). These are the only two localities that contain a mixture of haplotypes from two clades (Northwestern and Central). The first locality (19 in Figure 1) is located in a region that was previously presumed to be the location of a distinct subspecies, *M. p. kincaidi* (Hoffmann & Koeppl 1985). The second locality is located along a "suture zone", a region were secondary contact is common among other organisms (Swenson & Howard 2005). The location of contact between the Central and Eastern clades is currently unknown but may be elucidated with more extensive sampling north and south of the Great Lakes region.

Substructure within the Northwestern, Central, and Eastern clade is limited. Within the Northwestern clade, SE Alaska supported divergent populations on several islands; previous studies have shown that this region served as a Pleistocene Coastal refugium (Cook *et al.* 2001, 2006; Sawyer & Cook 2016; Figure 2). Admiralty Island supported a single haplotype from two localities. Farther south, Kadin Island and Mitkof Island also supported distinct populations. We found that the Central clade displayed two weakly supported groups, one located further north in the Rockies and into Canadian Provinces and the second located along the Southern Rockies from Montana to Mexico (Figure 1). Additionally, a population in northern New Mexico was distinct from all

surrounding areas. Lastly, the Eastern clade has genetic structure that is distributed throughout the East Coast sea board. Of note is the strong support a subgroup of populations that are located in Northern Quebec and Labrador. North Carolina, the southern extent of this clade, harbors another highly supported clade as well.

Species tree estimation and divergence dating

Multilocus species tree analyses produced strong support for all mtDNA clades (Figure 3). Initial analyses incorporated 1 mtDNA and 6 nuDNA loci (Figure 3); subsequently, we ran analyses with only nuDNA (Figure S2). Species tree analyses utilizing only nuDNA corroborated the strongly supported relationships between the Northwestern, Central and Eastern lineages, whereas the relationship of the Florida population was unstable. The most recent split, that between the Northwestern and Central clades, occurred around ~70 kya. This relationship is consistent with mtDNA phylogenetic analyses (Figure 2) where the only significantly supported relationship was between these two clades. The 2nd most recent split occurred around the timing of the Last Interglacial (LIG; ~125 kya), and represents a resolved relationship that received low support in mtDNA analyses. Finally, the oldest split occurred around ~250 kya during the Pre-Illinoian. The TMRCA for the *M. pennsylvanicus* lineages and *M. longicaudus* (~530 kya) is consistent with the date reported in Sawyer and Cook (2016).

Genetic diversity and Demographic History

Cytochrome b exhibits high genetic diversity (Table 3), with sequence distance between major clades ranging from 0.018 to 0.045. The haplotypes of the Northwestern and Central clade were least divergent from each other, forming the only well supported clade relationship within *M. pennsylvanicus* (Figure 2). The Northwestern clade and

Central clade both exhibited the highest diversity from the Florida clade. Often, clade comparisons resulted in the Florida clade exhibiting the highest genetic diversity. Nuclear loci divergence was much lower than that exhibited in Cytb, ranging from 0.002 to 0.016 (Appendix B). The nuclear gene *LCAT* was the least informative of all loci while *ETS2* sequences were the most informative.

Generally, *M. pennsylvanicus* is characterized by a dynamic history consisting of isolation that was subsequently followed by expansion. These clades are geographically structured and exhibit high genetic variation. Variation in Cytb was high for *M. pennsylvanicus*, with high haplotype diversity, nucleotide diversity, and average number of nucleotide differences (Table 2). At the clade level, each maintained high haplotype diversity but inversely a low nucleotide diversity and average number of nucleotide differences. The Northwestern clade exhibited the lowest average number of nucleotide differences ($k = 5.13$) and nucleotide diversity ($\pi = 0.005$) while the Eastern clade had the most variation ($k = 9.15$ and $\pi = 0.0087$). When clades were further partitioned by presumed glacial and ice free regions during the Pleistocene, most values remained comparable to clade-levels values. Two aspects of note pertain to the Central and Eastern glaciated regions. Within the glaciated region of the Central region there was a significant decrease in average number of nucleotide differences and nucleotide diversity. The Eastern clade demonstrated higher genetic diversity in both population indices despite it being in an area that was presumed to be covered by the Laurentide ice sheet while the southern region remained ice free during the Pleistocene.

Population statistics for *M. pennsylvanicus* only showed significant evidence for expansion based on F_s values, R2 and *D* values were both non-significant; however, all

clades displayed significant values consistent with a scenario of recent expansion. In general, glaciated regions for the Northwestern and Central regions showed significant signal indicative of population expansion, while areas that are presumed to have remained free of ice exhibited stability based on Fu's F_s, R2 and Tajima's D tests. The Eastern clade presented a conundrum. Contrary to a scenario of recent expansion that is supported in the other two clades, the glaciated region of the Eastern clade shows signals of a stable population with non-significant population statistics, which suggests this region may have harbored a refugium propagating genetic diversity. Further illustrating the monophyletic grouping and geographic structure within *M. pennsylvanicus*, the AMOVA indicates the majority of variance, 77.73%, is between the four clades and 22.37% is partitioned within populations (Table 3).

To test the scenario of recent expansion, we used mismatch distribution analyses, and found that *M. pennsylvanicus* exhibited a trimodal distribution (A; Figure 3). When clades were split between unglaciated and glaciated regions, all regions presumed to be ice free were multimodal along with the northern region of the eastern clade, indicative of stationary populations (Harpending *et al.* 1998). The Northwestern clade and glaciated region of the Central clade showed a unimodal distribution of pairwise differences, indicating that these unglaciated regions may have experiences a demographic expansion in the past (Rodgers & Harpending 1992). To investigate where the propagation of nucleotide differences may have originated, we further restricted mismatch distributions analyses in the Northwestern clade to individuals located in SE Alaska and for the Eastern clade to individuals in the highly supported subclade that included individuals from Quebec and Labrador.

To increase our understanding of the demographic expansion dynamics within *M. pennsylvanicus* clades, we used Extended Bayesian Skyline Plots (EBSPs) to estimate the timing of population growth. Analyses were calibrated with a Cytb molecular clock rate (0.072 substitutions/site/Myr, Hope *et al*. 2014) due to the lack of reliable fossil dating. The Northwestern clade entered a period of population growth approximately 50 KYA and has gradually increased through time to the present (Figure 6). The Central and Eastern clades exhibit population growth that precedes that seen in the Northwestern clade, beginning approximately 70 KYA.

Ecological Niche Modeling

The modern day predictions for Ecological Niche Models (ENMs) are largely in agreement with known distributions of *M. pennsylvanicus* (Figure 6). Clade specific models indicate that regions of occupation do vary throughout North America with overlap occurring between the three broadly distributed clades. The region of niche overlap between the Central and Eastern clades (Figure 6) was not extensively sampled and in this current study no localities exhibited mixed mtDNA haplotypes. The region of niche overlap between the Northwestern and Central clades is more extensive, and two localities (19 and 21; Figure 1) exhibit localities with mixed haplotypes. Additionally, it is of note that *M. pennsylvanicus* persist beyond what is the ideal environment. Within the Northwestern clade, localities 24 and 40, and within the Eastern clade, localities 53 and 62 are far outside of the niche that characterizes the remainder of the clade. To the south, locality 28 is found south of estimated distribution of the Central clade; however, this population is now extirpated (List *et al.* 2010) In the Southwest, ENMs do predict that there is more suitable habitat than that occupied by *M. pennsylvanicus*. This may be

due to a couple reasons. One possible explanation for this is the loss of mesic meadows due to a combination of climate change and anthropogenic habitat modification. Recent population extirpations have been attributed to the loss of suitable habitat due to anthropogenic change (List *et al.* 2010) and warming temperatures have decreased the range of the meadow vole (Anderson 1972).

The ENMs for the LGM predict that the range for *M. pennsylvanicus* if more restricted than modern distributions (Figure 6), this is to be expected as much of the current distribution was covered by ice sheets. Distribution overlap between LGM and modern distributions is minimal, only occurring in parts of the southern Rockies, Florida and along part of the southern East Coast Sea Board. LGM predictions do identify regions along the Mexico border and Texas where fossil evidence supports the past presence of *M. pennsylvanicus* in regions no longer occupied (Kurten and Anderson, 1980; Harris 2014). An extensive portion of the western US is supported as suitable habitat for *M. pennsylvanicus*; however, no fossil evidence or modern distribution would indicate *M. pennsylvanicus* occupied these areas. A large portion of these regions is occupied by closely related species (*M. canicaudus*, *M. townsendii, M. montanus* and *M. longicaudus*), possibly explaining the absence of *M. pennsylvanicus* in these regions. Furthermore, although ENM supports Beringia as a suitable refugium for *M. pennsylvanicus*, our current genetic data does not support this refugium as a likely source for population expansion. This may be partially due to limited sampling and would benefit from more comprehensive sampling in this region.

Discussion

North American species with transcontinental distributions provide a unique opportunity to better understand what role historical biogeographic events of the late Pleistocene played in modern genetic structure and diversity (Dragoo *et al.* 2006; Burbrink *et al.* 2008; Puckett *et al.* 2015). Key benefits of studying the demographic and phylogeographic history of wide ranging species is that it permits assessment of barriers to gene flow, refines understanding of known and cryptic refugia, identifies cryptic species, and sheds light on biogeographic history of the region. Multiple divergent clades that are geographically structured are commonly found among many organisms in North America despite apparently continuous distributions (Reding *et al.* 2012; Puckett *et al.* 2015; Hope 2016). However, both the location of boundaries and the depth of divergence between clades is not always the same, highlighting the idiosyncratic nature of species response to changing environments (Webb 1988; Soltis 2006).

Our data and analyses indicate that the cyclic climatic fluctuation of the Late Pleistocene played an important role in the ecology and evolution of *M. pennsylvanicus.* By integrating multilocus genetic analyses and environmental informatics we detect significant phylogeographic signatures with implications for at least three aspects of the history of *M. pennsylvanicus*. First, the biogeographic history appears to be more complex that previously considered. Second, the post-glacial (re)colonization and demographic signatures suggest recent expansion into northern latitudes. Third, we find evidence for at least four areas likely occupied during the last glacial period (SE Alaska, Southwest, Southeast and Florida). For example, the multilocus genetic diversity corresponds with well-predicted geographic distribution models resulting in phylogeographic structure. Moreover, we documented that the timing of climatic changes

dictated the timing and mode of major demographic events. When combined, the genetic patterns and ENMs reflect that *M. pennsylvanicus* experienced a history of range contraction to refugia during glacial and interglacial periods with subsequent expansion into previously glaciated areas following the retreat of the Laurentide and Cordilleran ice sheets.

Additionally, ENMs support that each clade, based on molecular data, occupies a separate niche with overlap in regions where these two clades have come into contact since the LGM. Through range-wide sampling and phylogenetic analyses we confirmed the presence of 3 broadly distributed clades and one spatially restricted population in Florida (Jackson 2016). The clades are largely parapatric with only two localities (19, 20; Figure 1) exhibiting mixed haplotypes from the two westernmost clades (Northwestern and Central), indicative of secondary contact. Finally, mtDNA provides evidence that the clades are paraphyletic in regards to the nominal species *M. breweri*.

Biogeographic history of Microtus pennsylvanicus

The deepest split in *M. pennsylvanicus* is between the Eastern clade and Florida population. Our species tree analyses date the split to the pre-Illinoian ~250 kya (Figure 3); however, the relationship of Florida to the other *M. pennsylvanicus* clades remains uncertain due to low posterior branch support in our Cytb and species tree analyses (Figure 1 and 2). Genetically, this population is highly divergent (>3.8%, Table 3) from all other clades and all individuals sequenced shared a single Cytb haplotype, despite being caught in different localities 32 years apart. High genetic divergence but non-existent intraclade variation may be attributable to the dynamic biogeographic history of the Florida Peninsula. In the time since the pre-Illinoian, exposed land in the Florida

Peninsula would have expanded and contracted with sea level fluctuations (Webb, 1990) and at various intervals would have been separated from the continental US. During periods of inundation, peninsular populations would have contracted to a series of ridges in central Florida, essentially surviving as island populations (Webb 1990; Clark *et al.* 1999). The fossil record supports *M. pennsylvanicus* as a mammal that has persisted in Florida for an extended period of time; however, our species tree analyses indicate that isolation predates the available fossil record (Kurten & Anderson 1980; Webb and Wilkins 1984). ENMs support an extensive region along the Gulf Coast as suitable habitat for *M. pennsylvanicus* during the LGM. Furthermore, the phylogeographic break between the Atlantic Coast and Gulf Coast has been reported for other organisms (Soltis 2006) and coincides with high divergence between the Florida population and Eastern Clade. Our study supports the Florida population as a genetically and ecologically distinct lineage based on molecular analyses and ENMs, a result that contributes to the recognition of a growing number of genetically distinctive populations of vertebrates that are endemic to the Florida Peninsula including the beach mouse (*Peromyscus polionotus*; Avise *et al.* 1983), the North American racer (*Coluber constrictor*; Burbrink *et al.* 2008) and the mud turtle (*Kinosternon bauri*; Walker *et al.* 1998).

The next major division within *M. pennsylvanicus* is estimated to have occurred ~125 kya (Figure 3) between the Eastern clade and the two westernmost (Central and Northwestern; Figure 1) and is congruent with the Last Interglacial (LIG, ~120-140 kya). This east-west break is a common zoogeographic pattern that is mirrored in other North American species such as *Myodes gapperi* (Runck & Cook 2005); *Ursus americanus* (Wooding & Ward 1997; Stone & Cook 2000), and was earlier identified in populations

of *M. pennsylvanicus* in Canada using restricted fragment length polymorphisms (Plante *et al.*, 1989). Furthermore, the genetic distance between the Eastern and western clades (4.2% and 3.9% for the Northwestern and Central clades respectively; Table 3) is comparable to that in other taxa (e.g., Runck & Cook 2005). Hall & Kelson (1959) speculated this common split may be due to the aridification of the mid-continental region during Pleistocene interglacials. During these periods, forest specialist and mesic-associated organisms would have persisted in eastern and western refugia. Subsequently following the interglacials, organisms from both of these refugia expanded and contacted in the Great Plains region (Swenson & Howard 2004).

The final division between the Central and Northwestern clades is the most recent. Species tree analyses estimate that the split occurred before the Last Glacial Maximum ~70 kya (Figure 3). In our Cytb analyses, this is the only relationship strongly supported within *M. pennsylvanicus* clades (Figure 2). Additionally, ENMs support a wide range of niche overlap where secondary contact has occurred between these two clades at two localities (Figure 6). Uncorrected genetic divergence between the Northwestern and Central clades is 1.8% based on Cytb, which is comparable to the genetic distance between the *M. breweri* and the Eastern clade (1.8%, Table 3). Based on the lack of genetic structure throughout mainland Alaska and northwest Canada, we infer that *M. pennsylvanicus* persisted in coastal refugia near the Alexander Archipelago (AA) of SE Alaska. We presume it is unlikely that *M. pennsylvanicus* was in the Beringian Refugium due to no fossil record and low genetic structure and low diversity in contemporary populations in Interior Alaska (Hewitt, 2000). Alternatively, studies have proposed that the AA served as a coastal refugium for other mammalian species (e.g.

Heaton *et al.*, 1996; Fleming and Cook, 2002; Lucid and Cook, 2004; Sawyer and Cook, 2016). Refugial populations led to paleoendemic taxa that are genetically distinct from mainland populations (Dawson *et al.*, 2007). Within our mtDNA analyses, the only regions that retained any structure were insular populations in SE Alaska (Figure 2). We consistently found strong support for a group that consisted of individuals from Admiralty Island and another that consisted of Mitkof Island, Kadin Island and a mainland locality, suggesting that this group has expanded from the islands to mainland.

Postglacial Colonization and Demographic Analyses

To investigate the postglacial dynamics within *M. pennsylvanicus,* we explored summary and demographic statistics at multiple hierarchical levels: first, *M. pennsylvanicus* as a whole; second, at the clade level; and finally we examined each clade based on the proximity of specimen localities to glacial extent during the LGM. Cytb data for *M. pennsylvanicus* are characterized by high haplotype diversity (*Hd*), nucleotide diversity (π) and average number of pairwise differences (*k*; Table 2). Generally, each widely distributed clade was characterized by signals of recent population expansion and deep genetic divergence. Summary statistics calculated for each clade (Fu's Fs, R2, and Tajima's D) were all significant, corroborating the scenario of recent range expansion within each. Eastern and Central clades both retained high *Hd* and a high *k*, conversely, π and *k* were low in the Northwestern clade (Table 2), indicating a strong signal of recent expansion into Alaska and Canada.

Populations that are now in deglaciated regions should show signatures of recent expansion from refugia, while southern populations should have higher genetic diversity and exhibit signatures of longterm persistence (Hewitt 2000; Lessa *et al.* 2003). We

expected to see these signatures in comparisons between northern glaciated and southern unglaciated localities in *M. pennsylvanicus* and post-glacial expansion in the meadow vole largely coincides with the generalized patterns of colonization in North America (Arbogast 1999; Runck & Cook, 2005; Dragoo *et al.* 2006; Malaney & Cook 2013, Sawyer *et al.* 2016), with a few exceptions. Following retreat of the ice sheets, organisms tracked newly available environments into the tabula rasa of the higher latitudes (Hewitt 2000). Summary statistics significantly supported recent demographic expansion for populations located in deglaciated regions within the Northwestern and Central clades. During the LGM (Figure 7) these regions were covered by the Cordilleran and Laurentide ice sheets, with demographic analyses revealing high *Hd*, low π, (Table 2) and mismatch distribution plots that were strongly unimodal (Figure 3). All these metrics are consistent with rapid expansion as new habitat became available. Presumed refugia (southern end of the Rockies in the Southwest for the Central clade and SE Alaska for the Northwestern clade) for these colonization events are both supported as highly suitable for *M. pennsylvanicus* during the LGM based on the historical ENMs (Figure 7). Summary statistics for these purported refugia are now consistent with recent colonization of these regions (Table 2).

In the presumed glacial refugium for the Central clade near the southern terminus of the Rocky Mountains, we document expected patterns of prolonged persistence that is characterized by deep divergence, high nucleotide and haplotype diversity and genetic structure (e.g., mtDNA analyses show strong support for multiple divergent clades in New Mexico). Previous studies focusing on jumping mice (Malaney & Cook 2013) also illustrated that this region, near the southern terminus of the Rocky Mountains, also

harbored high genetic structure. Within *M. pennsylvanicus*, individuals in the Santa Fe National Forest were >1.5% from nearby localities in NM and Colorado. Additionally, the now extirpated population in Chihuahua was >1% divergent from all other populations in the Southern Rockies, suggesting that the high genetic diversity and structure in the Southwest is consistent with the hypothesis that *M. pennsylvanicus* persisted in this region during the Pleistocene (e.g., Malaney & Cook 2013). We suggest that the Central clade of *M. pennsylvanicus* expanded north from the Southwest (southern limit of the Rocky Mountains), based on genetic and demographic analyses and fossil evidence. These findings are consistent with previous phylogeographic studies that found populations of a number of organisms persisted through the LGM in refugia east of the Rocky Mountains and south of the ice sheets and subsequently moved northward as boreal vegetation recolonized the high latitudes (Jaramillo-Correa *et al.* 2004; Runck & Cook, 2005; Hope *et al*. 2012; Dawson *et al.* 2014).

Cryptic Refugia in Northeastern Canada

This study raises the possibility of a refugium in northeastern Canada during the LGM that impacted genetic structure in eastern populations of *M. pennsylvanicus.* Demographic analyses also support the hypothesis that there were eastern populations of meadow voles that persisted through the LGM south of the Laurentide ice sheet (Table 2, Figure 3) essentially creating zone of contact in southern Quebec where voles from two independent sources are now in contact. In contrast to other eastern species, the dynamics of post-Pleistocene colonization of northeastern Canada by meadow voles was more complex than in other mammals such black bears (Paetkau & Strobeck, 1996; Marshall *et al*. 2011), woodland caribou (Wilkerson, 2010) and red foxes (Langille *et al.* 2014).

Wilkerson (2010) referred to two possible routes of colonization of Newfoundland as the Northern and Southern routes (Figure 8). In the northern route scenario, organisms originating in the southeastern refugium (south of the ice sheets) followed the contraction of Laurentide ice sheet northward into Quebec and Labrador and colonized Newfoundland across the Strait of Belle Isle (Figure 8). In this case, populations on Newfoundland would be closely related to those found in Quebec and Labrador, all of which would be descendants from colonizers originating south of the ice sheets. The southern route scenario involves populations from either island refugia or the eastern coastal plains along the east coast of North America (Figure 8; Pielou 1991). In this case, a close relationship be seen between the populations on Newfoundland and those in New England and Canadian provinces along the Atlantic (e.g., Nova Scotia). Previous studies (e.g., Wilkerson 2010; Langille *et al.* 2014) have shown that larger mammalian fauna tracked the contraction of the ice sheets and populated Newfoundland through the northern route.

Our data suggest instead that populations in Quebec and Labrador may have been isolated for an extended period of time and that it is unlikely that the colonization of Newfoundland occurred over the Strait of Belle Isle. In our Cytb analyses there is strong support for a basal group that consists only of individuals in Northern Quebec and Labrador. These populations, in contrast to previous studies, do not share a close relationship with populations on Newfoundland, as expected if expansion occurred in the same manner as larger mammals (Wilkerson 2010; Langille *et al.* 2014). Additionally, summary statistics (Fu's Fs, R2, and Tajima's D; Table 3) do not significantly support a scenario of recent expansion for the previously glaciated region of NE Canada. Average

pairwise differences were high for NE Canada (k = 9.20, Table 3), higher than that seen in the previously glaciated regions in the central or northwestern clades. Generally, high *Hd* and high π are signals of stability for a population; conversely, *k* is indicative of a recent expansion. In the case of NE Canada, *Hd* and π values are comparable to other glacial refugia (*Hd* = 0.971 and π = 0.00836) within *M. pennsylvanicus*. To further investigate the scenario of long-term persistence in NW Canada, we calculated summary statistics for populations between two presumed refugia (populations south of locality 60 and north of 58 in the Eastern clade); however, we did not find any signal of recent expansion. One locality (57; Figure 1) did harbor mixed haplotypes from the NW Canada subclade and southern subclade, which may indicate this region is an area of contact between different refugial populations. Two possible explanations for the lack of a signal of recent expansion are: 1) two independent refugial sources are now coming into secondary contact and these two populations are exhibiting long term persistence despite recent expansion, or 2) high nucleotide diversity and no signature of expansion may be the result of gradual colonization, where the slower expansion retained high levels of genetic diversity (i.e., phalanx model; Nichols & Hewitt 1994). In either case, our results are indicative of a persistent glacial refugium in NE Canada. Future studies should utilize more comprehensive sampling to better characterize the postglacial demographic dynamics of meadow voles and other species in this region.

Conclusion

The complex and dynamic biogeographic history of the meadow vole illustrates the key role that climatic fluctuations had on structuring contemporary genetic diversity. Additionally, our findings suggest the presence of cryptic refugia in northeastern Canada.

These findings are consistent with previous studies, noting the idiosyncratic nature of species response to Late Pleistocene environmental cycling. Despite occupation in similar refugia, how species responded to climate warming and the retreat of ice sheets following the LGM varied. *Microtus pennsylvanicus* is characterized by diversification through isolation in separate refugia subsequently followed by rapid expansion with secondary contact between distinct lineages. Furthermore, our ENMs suggest that the three geographically distributed clades inhabit environments with minimal spatial overlap and the Florida population is restricted to the south. Additionally, ENMs predict that the range of *M. pennsylvanicus* was previously more extensive throughout the southeastern US than today, causing further concern for the persistence of *M. dukecampbelli*. Through comprehensive phylogeographic studies we gain a better understanding of how historical climate change has affected species distributions, this in turn, allows us to better manage and predict how organisms will respond to modern climate change.

References

Anderson, R.P., & Gonzalez I.J. (2011) Species-specific tuning increases robustness to sampling bias in model of species distributions: An implementation with Maxent. *Ecological Modelling*, **222**, 2796-2811.

Anderson, R.P., & Raza A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378-1393.

Anderson, S., & Hubbard J.P. (1971) Notes on Geographic Variation of *Microtus pennsylvanicus* (Mammalia, Rodentia) in New Mexico and Chihuahua. *American Museum Novitates*, **2460**, 1-8.

Araujo, M.B., & Peterson A.T. (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology*, **93**, 1527-1539.

Arbogast, B.S. (1999) Mitochondrial DNA phylogeography of the new world flying

squirrels (*Glaucomys*): Implications for Pleistocene biogeography. *Journal of Mammalogy*, **80**, 142-155.

Avise, J.C., Shapira, J.F., Daniel, S.W., Aquadro, C.F., & Lansman, R.A. (1983) Mitochondrial DNA differentiation during the speciation process in *Peromyscus. Molecular Biology and Evolution*, **1**, 38-56.

Avise, J.C. (2000) *Phylogeography: the history and formation of species*. Harvard Univeristy. Press, Cambridge, MA, USA.

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

Barbet-Massin, M., Jiguet F., Albert, C.H., & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327-338.

Barve, N., Barve, V., Jimenex-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberon, S.T., & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810-1819.

Brown. J.L. (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, **5,** 694-700.

Brunsfeld, S.J., Sullivan, J., Soltis, D.E., & Soltis, P.S. (2001) Comparative phylogeography of northwestern North America: a synthesis. Integrating Ecological and Evolutionary Processes in a spatial context (eds. J. Silvertown, and J. Antonovics), pp. 319-339. Blackwell Science, Oxford.

Burbrink, F.T., Fontanella, F., Pyron, R.A., Guiher, T.J., & Jimenez, C. (2007) Phylogeography across a continent: The evolutionary and demographic history of the North America racer (Serpentes: Colubridae: *Coluber constrictor*). *Molecular Phylogenetics and Evolution*, **47**, 274-288.

Anderson, D.R., & Burnham, K.P. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*. 2nd ed. Springer-Verlag. New York, NY, USA.

Clark, A.M., Bowen, B.W. & Branch, L.C. (1999) Effects of habitat fragmentation on an endemic scrub lizard (*Sceloporus woodi*): an historical perspective based on a mitochondrial gene genealogy. *Molecular Ecology*, **8,** 1093-1104.

Conroy, C.J., & Cook, J.A. (2000) Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Rodentia: Muridae). *Molecular Ecology*, **9**, 165-175.

Conroy C.J., & Neuwald, J.L. (2008) Phylogeographic study of the California vole, *Microtus californicus*. *Journal of Mammalogy*, **89**, 755-767.

Cook, J.A., Bidlack, A.L, Conroy, C.J., Demboski, J.R., Fleming, M.A., Runck, A.M., Stone, K.D., & MacDonald, S.O. (2001) A phylogeographic perspective on endemism in the Alexander Archipelago of southeast Alaska. *Biological Conservation*, **97**, 215-227.

Cook, J.A., Dawson, N.G., & MacDonald, S.O. (2006) Conservation of highly fragmented systems: the north temperate Alexander Archipelago. *Biological Conservation*, **133**, 1-15.

Crawford, D.L., Dragoo, J.W., Smith, F.A., & Chavez, A.N. (2011) Diversification within the Mexican Vole (*Microtus mexicanus*) and the Role of Post-Pleistocene Climate Change. *Western North American Naturalist*, **71**, 176-194.

Darriba D, Taboada, G.L., Doallo, R., & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Dawson, N. G., MacDonald, S.O. & Cook, J.A. (2007) Endemic Mammals of the Alexander Archipelago. Wildlife management of the Tongass National Forest (ed. By J. Schoen). Audubon Special Publication, New York.

Dawson, N.G., Hope, A.G., Talbot, S.L. & Cook, J.A. (2014) A multilocus evaluation of ermine (*Mustela ermine*) across the Holarctic, testing hypothesis of Pleistocene diversification in response to climate change. *Journal of Biogeography*, **41**, 464-475.

DeChaine, E.G. (2008) A bridge or a barrier? Beringia's influence on the distribution and diversity of tundra plants. *Plant Ecology & Diversity*, **1**, 197-207.

Demboski, J.R., & Cook, J.A. (2001) Phylogeography of the dusky shrew, *Sorex monticolus*, (Insectivora, Soricidae): insight into deep and shallow history in northwestern North America. *Molecular Ecology*, **10,** 1227-1240.

DeWoody, J.A., Chesser, R.K, & Baker, R.J. (1999) A translocated mitochondrial cytochrome *b* pseudogene in voles (Rodentia: *Microtus*). *Journal of Molecular Evolution*, **48**, 380-382.

Dragoo, J. W., Lackey, J.A., Moore, K.E., Lessa, E.P., Cook, J.A. & Yates, T.L. (2006) Phylogeography of the deer mouse *(Peromyscus maniculatus)* provides a predictive framework for research on hantaviruses. *Journal of General Virology*, **87**, 1997- 2003.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A. (2012) Bayesian phylogenetics

with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969-1973.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792-1797.

Elith, J. *et al.*, (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.

Elith, J., Phillips, S.J., Hastie, T., Dudik, M., Chee, Y.E, & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43-57.

Ellsworth, D.L., R.L. Honeycutt, N.J. Silvy, J.W. Bickham and W.D. Klimstra. 1994. Historical Biogeography and Contemporary Patterns of Mitochondrial DNA Variation in White-Tailed Deer from the Southeastern United States. Evolution, 48:122-136.

Fedorov, V.B., Goropashnaya, A.V, Boeskorov, G.G., & Cook, J.A. (2007) Comparative phylogeography and demographic history of the wood lemming (*Myopus schisticolor*): implications for late Quaternary history of the taiga species in Eurasia. *Molecular Ecology*, **17**, 598-610.

Fourcade, Y., Engler, J.O., Rodder, D., & Secondi, J. (2014) Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS ONE*, 9:e97122.

Fu, Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

Garroway, C.J., Bowman, J., Cascaden, T.J., Holloway, G.L., Mahan, C.G., Malcolm, J.R., Steele, M.A., Turner, G., & Wilson, P.J. (2009) Climate change induced hybridization in flying squirrels. *Global Change Biology*, **16**,113-121.

Guiller-Arroita, G., *et al.* (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, **24**, 276-292.

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

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

Haring, E., Sheremetyeva, I.N., & Kryukov, A.P. (2011) Phylogeny of Palearctic vole

species (genus *Microtus*, Rodentia) based on mitochondrial sequences. *Mammalian Biology*, **76**, 258-267.

Harpending, H.C., Batzer, M.A., Gurvens, M., Jorde, L.B., Rogers, A.R., & Sherry S.T. (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences*, **95**, 1961-1967.

Helend, J., & Drummond, A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.

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

Hewitt, G.M. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.

Hewitt, G.M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **359**, 183-195.

Hoatling, A.S., Percival, H.F., Kitchens, W.M., & Kasbohm, J.W. (2010) The Persistence of Endangered Florida Salt Marsh Voles in Salt Marshes of the Central Florida Gulf Coast. *Southeastern Naturalist*, **9**, 795-802.

Hoffmann, R.S., & Koeppl, J.W. (1985) Zoogeography. Biology of New World *Microtus*. (eds. Tamarin, R.H) pp. 85-115. Special Publication, American Society of Mammalogist.

Hofreiter, M., & Stewart, J. (2009) Ecological change, range fluctuations and population dynamics during the Pleistocene. *Current Biology*, **19**: 584-594.

Hope, A.G., Speer, K.A, Demboski, J.R., Talbot, S.L., & Cook, J.A. (2012) A climate for speciation: rapid spatial diversification within the *Sorex cinereus* complex of shrews. *Molecular Phylogenetics and Evolution*, **64**, 671-684.

Hope, A.G., Panter, N., Cook, J.A., Talbot, S.L., & Nagorsen, D.W. (2013) Multilocus phylogeography and systematic revision of North American water shrews (genus: *Sorex*). *Journal of Mammalogy*, **95**, 722-738.

Hope, A.G., Ho, S.Y.W., Malaney, J.L., Cook, J.A., & Talbot, S.L. (2014) Account for rate variation among lineages in comparative demographic analyses. *Evolution*, **68**, 2689-2700.

Hope, A.G., Waltari, E., Malaney, J.L., Payer, D.C., Cook, J.A., & Talbot, S.L. (2015) Arctic biodiversity: increasing richness accompanies shrinking refugia for a cold-associated tundra fauna. *Ecosphere*, **6**, 1-67.

Hubbard, J.P., Thaeler Jr., C.S., & Schmitt, C.G. (1983) Notes on voles (*Microtus*, Family Cricetidae) in New Mexico. *Occasional Papers of the Delaware Museum of Natural History*, **28**, 1-9.

Huelsenbeck, J.P. & Ronquist, R. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754-755.

Irwin, D. M., Kocher, T. D., *et al*. (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution*, **32**, 128-144.

Jaramillo-Correa, J.P., Beaulieu, J., & Bousquet, J. (2004) Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (*Picea mariana*) a transcontinental North American conifer. *Molecular Ecology*, **13**, 2735-2747.

Kearse, M., Moir, R., Wilson, A., *et al*. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647-1649.

Kumar, S., Stecher, G., & Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*. Doi: 10.1093/molbev/msw054.

Kurten, B., & Anderson, E. (1980) Pleistocene mammals of North America. Columbia University Press, New York, NY, USA.

Langille, B.L., O'Leary, K.E., Whitney, H.G., & Marshall, H.D. (2014) Mitochondrial DNA diversity and phylogeography of insular Newfoundland red foxes (*Vulpes vulpes deletrix*). *Journal of Mammalogy*, **95**, 772-780.

Lessa, E.P., Cook, J.A., & Patton, J.L. (2003) Genetic footprints of demographic expansion in North America, but no Amazonia, during the Late Quaternary. *Proceedings of the National Academy of Sciences*, **100**, 10331-10334.

Librado, P., & Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.

List, R., Pergams, O.R.W., Pachecho, J., Cruzado, J., & Ceballos, G. (2010) Genetic divergence of *Microtus pennsylvanicus chihuahuensis* and conservation implications of marginal population extinctions. *Journal of Mammalogy*, **91**, 1093-1101.

Malaney, J.L., & Cook, J.A. (2013) Using biogeographical history to inform conservation: the case of Preble's meadow jumping mouse. *Molecular Ecology*, **22**, 6000-6017.

Marshall, H.D., Yaskowiak, E.S., Dyke, C., & Perry, E.A. (2011) Microsatellite population structure of Newfoundland black bears (*Ursus americanus hamiltoni*). *Canadian Journal of Zoology*, **89**, 832-839.

Martinkova N., and Moravec, J. (2012) Multilocus phylogeny of arvicoline voles (Arvicolini, Rodentia) shows small tree terrace size. *Folia Zoology*, **61**, 254-267.

McCormack, J.E., Zellmer, A.J., & Knowles, L.L. (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: Insights from tests with niche models. *Evolution*, **64**, 1231-1244.

Merow, C., Smith, M.J., & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, **36**, 1058-1069.

Nichols, R.A., & Hewitt, G.M. (1994) The genetic consequences of long distance dispersal during colonization. *Heredity*, **72**, 312-317.

Paetkau, D., & Strobeck, C. (1996) Mitochondrial DNA and the phylogeography of Newfoundland black bears. *Canadian Journal of Zoology*, **74**, 192-196.

Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637-669.

Pearson, R.G., Raxworthy, C.J., Nakmura, M., & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102-117.

Peterson, A.T., Papes, M., & Eaton, M. (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography*, **30**, 550-560.

Phillips, S.J., & Dudik, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161-175.

Pielou, E.C. (1991) After the Ice Age. University of Chicago Press, Chicago, IL, USA.

Plante, Y., Boag, P.T., & White, B.N. (1989) Macrogeographic variation in mitochondrial DNA of meadow voles (*Microtus pennsylvanicus*). *Canadian Journal of Zoology*, **67**, 158-167.

Platt, A.R., Woodhall, R.W., & George, Jr, A.L. (2007) Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. *BioTechniques*, **43**, 58-62.

Puckett, E.E., Etter, P.D., Johnson, E.A., & Eggert, L.S. (2015) Phylogeographic analyses of American black bears (*Ursus americanus*) suggest four glacial refugia and complex patterns of postglacial admixture. *Molecular Biology and Evolution*, **32**, 2338-2359.

Ramos-Onsins, S.E., & Rozas, J. (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092-2100.

Rambaut, A., & Drummond, A.J. (2007) *Tracer v1.4.* University of Edinburgh, Edinburgh, UK. Available at: http://beast.bio.ed.ac.uk/tracer.

Rambaut, A. (2009) FigTree. http://tree.bio.ed.ac.uk/software/figtree/

Reding, D.M., Bronikowski, A.M., Johnson, W.E. and Clark, W.R. (2012) Pleistocene and ecological effects on continental-scale genetic differentiation in the bobcat (*Lynx rufus*). *Molecular Ecology*, **21**, 3078-3093.

Richmond, O.M.W., McEntee, J.P., Humans, R.J., & Brashares, J.S. (2010) Is the climate right for pleistocene rewilding? Using species distribution models to extrapolate climatic suitability for mammals across continents. *PLos ONE*, 5:e12899.

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

Rogers, A.R, &Harpending, H.C. (1992) Population growth makes waves in the distribution of pairwise nucleotide differences. *Molecular Biology and Evolution*, **9**, 552-569.

Runck, A.M., & Cook, J.A. (2005) Postglacial expansion of the southern red-backed vole (*Clethrionomys gapperi*) in North America. *Molecular Ecology*, **14**, 1445-1456.

Runck, A.M., Matocq, M.D., & Cook, J.A. (2009) Historic hybridization and persistence of a novel mito-nuclear combination in red-backed voles (genus *Myodes*). *BMC Evolutionary Biology*, **9**, 114.

Sawyer, Y.E., & Cook, J.A. (2016) Phylogeographic structure in long-tailed voles (Rodentia: Arvicolinae) belies the complex Pleistocene history of isolation, divergence, and recolonization of Northwest North America's Fauna. *Ecology and Evolution*,

Shafer, A.B., Cullingham, C.I., Cote, S.D., & Coltman, D.W. (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589-4621.

Shcheglovitova, M., & Anderson, R.P. (2013) Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. *Ecological Modelling*, **269**, 9-17.

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

Soltis, P.S. (2003) Hybridization, speciation and novelty. *Journal of Evolutionary Biology*, doi: 10.1111/jeb.12095

Stephens, M., Smith, N.J. & Donnelly, P. (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978-989.

Stone, K.D., & Cook, J.A. (2000) Phylogeography of black bears (*Ursus americanus*) from the Pacific Northwest. *Canadian Journal of Zoology*, **78**, 1-6.

Stone, K.D., Flynn, R.W., & Cook, J.A. (2002) Postglacial colonization of northwestern North America by the forest-associated American marten (*Martes americana*, Mammalia: Carnivora: Mustelidae). *Molecular Ecology*, **11**, 2049-2063.

Swenson, N.G., and Howard, D.J. (2004) Do suture zones exist? *Evolution*, **58**, 2391-2397.

Swenson, N.G., & Howard, D.J. (2005) Clustering of Contact Zone, Hybrid Zones, and Phylogeographic Breaks in North America. *American Society of Naturalis*, **116**, 581-591.

Tajima, F. (1989) Statistical method for testing the Neutral Mutation Hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

Thuiller, W., Lafourcade, B., Engler, R., & Araujo, M.B. (2009) BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-373.

Tougard, C., Brunet-Lecomte, R., Fabre, M.,& Montuire, S. (2008) Evolutionary history of two allopatric *Terricola* species (Arvicolinae, Rodentia) from molecular, morphological, and palaeontological data. *Biological Journal of the Linnean Society*, **93**, 309-323.

Walker, D., Molar, P.E., Buhlmann, K.A. & J.C. Avise. (1998) Phylogeographic patterns in *Kinosternon subrubrum* and *K. bauri* based on mitochondrial DNA restriction analysis. *Herpetologica*, **54**, 174-184.1.

Waltari, E., Hijmans, R.J, Peterson, A.T., Nyari, A.S, Perkins, S.L., & Guralnick, R.P. (2007) Locating Pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLos ONE* 2:e563.

Waltari, E., and Guralnick, R.P. (2009) Ecological niche modelling of montane mammals in the Great Basin, North America: examining past and present connectivity of species across basins and ranges. *Journal of Biogeography*, **36**, 148-161.

Warren, D.L., R.E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of enironmental niche models. *Ecography*, **33**, 607-611.

Warren, D.L., & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335-342.

Warren, D.L., Wright, A.N., Seifert, S.N., & Shaffer, A.B. (2014) Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Diversity and Distributions*, **20**, 334-343.

Webb, S. D., & Wilkins, K.T. (1984) Historical biogeography of Florida Pleistocene mammals. Contributions in Quaternary vertebrate paleontology: a volume in memorial to John E. Guilday (eds. Genoways, H.H. & Dawson, M.R.) Pp. 370-383. Carnegie Museum of Natural History Special Publications 8.

Webb, S.D. (1990) Historical biogeography. Ecosystems of Florida (eds. Myers, R.L., & Ewel, J.J.) pp. 70-102. University Central Florida Press, Orlando.

Wilkerson, C.W. (2010) Population genetics of woodland caribou (*Rangifer tarandus caribou*) on the island of Newfoundland, M.S. thesis. Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada.

Wooding, S., & Ward, R. (1997) Phylogeography and Pleistocene evolution in the North American black bear. *Molecular Biology and Evolution*. **14**, 1096-1105.

Figures and Table

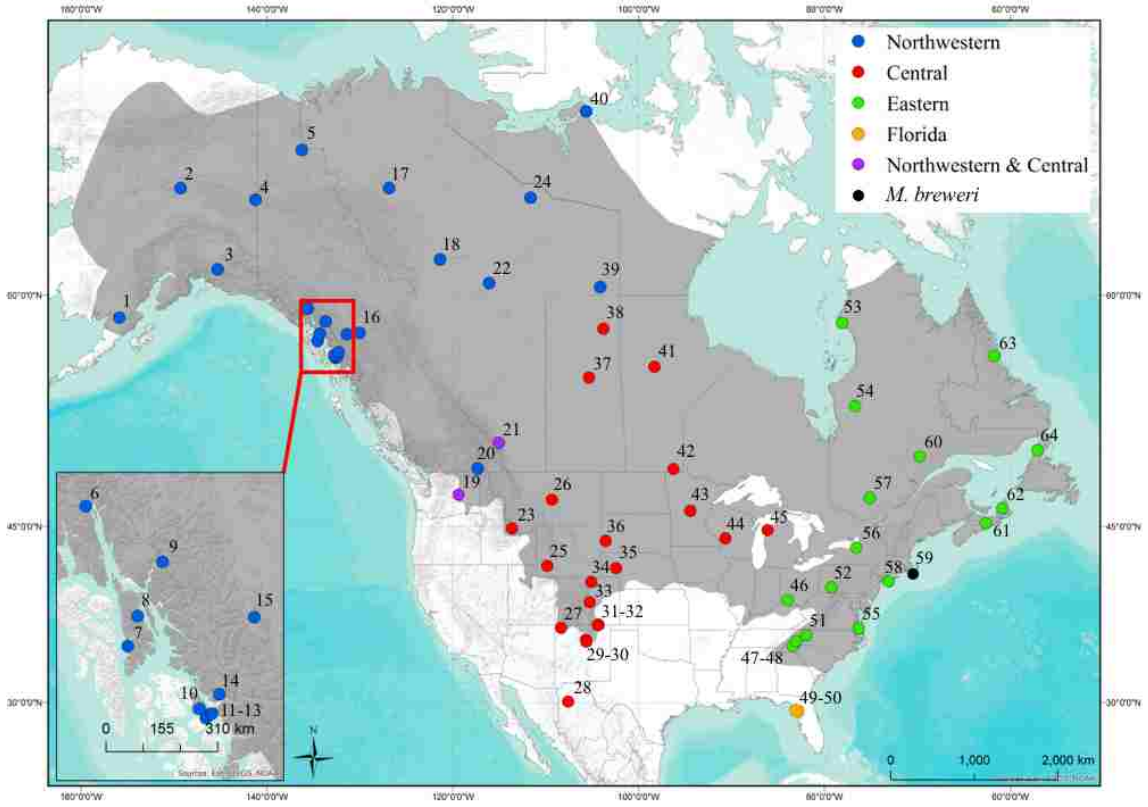


Figure 1. Distribution of *M. pennsylvanicus* is shown in gray (modified from IUCN) with dots representing sampling localities. Colors correspond to four distinct *Cyt b* clades, purple dots are localities with Northwestern and Central individuals, they do not represent a distinct clade. Localities are numbered sequentially from west to east.

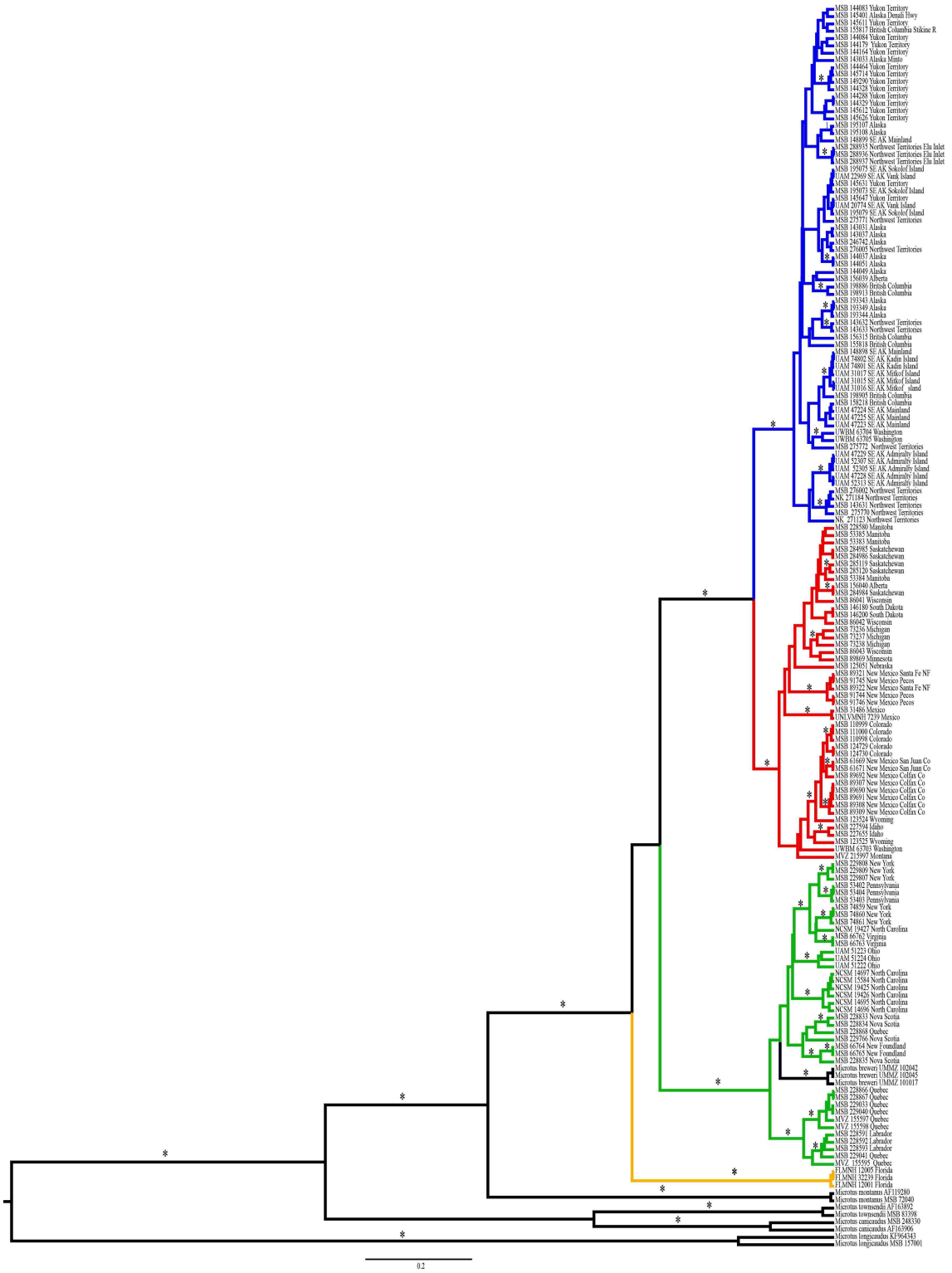


Figure 2. *Microtus pennsylvanicus* Cytb gene tree for 157 individuals (379-1140 bp). Colors correspond to major clades in Figure 1. Two individuals for each outgroup (*M. longicaudus, M. townsendii, M. canicaudus, M. montanus* and *M. breweri*). Posterior probability >0.95 is indicated by an asterisk.

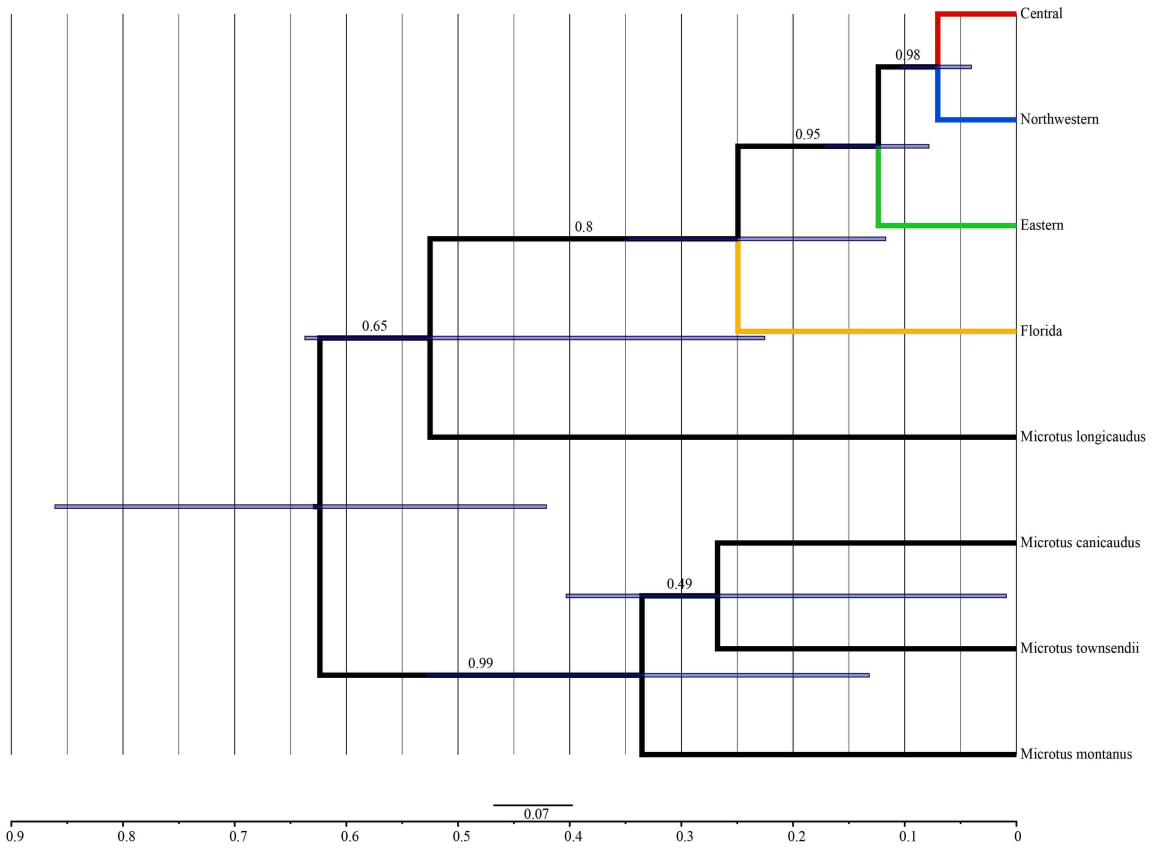


Figure 3. Species-tree estimation for *Microtus pennsylvanicus* clades and 4 closely related *Microtus* species based on one mitochondrial and six nuclear loci. Phylogeny estimation was conducted in *BEAST providing both divergence estimates (show in italics to the right) and posterior nodal support values (shown in bold). A timescale is included with present day to the right and past to the left in millions of years. Additionally, the 95% confidence intervals for the date estimates is shown (blue bars).

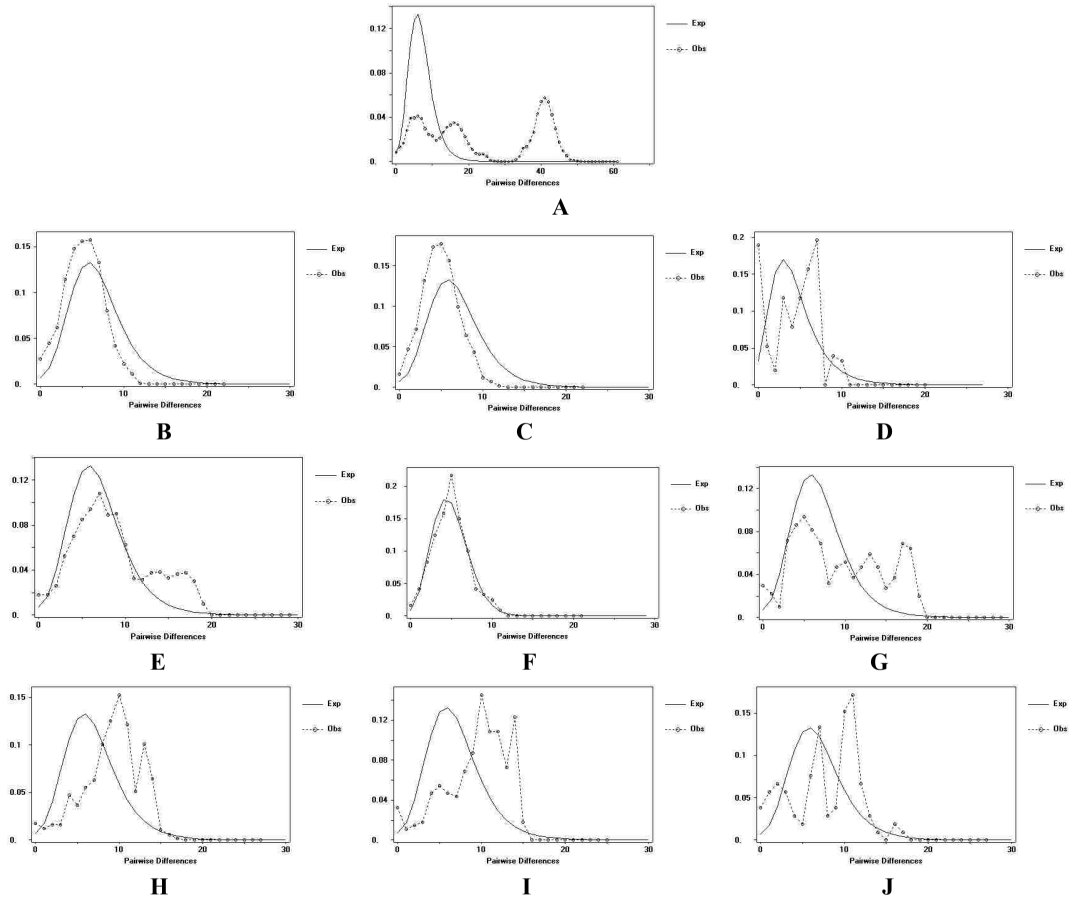


Figure 4. Mismatch distribution plots for all available samples (A, N = 145), and for each major lineage with the exception of the Florida samples (N = 3). Plots B-D are for the Northwestern lineage, E=G for the Central lineage and H-J for the Eastern lineage. The First column for each lineage exhibits the mismatch distribution of the clade as a whole, the second column only includes localities in previously glaciated regions and the final column represents regions that were not covered by ice sheets during the Pleistocene. Solid lines indicate the expected distribution under a model of sudden population expansion and a dashed line represents the observed frequency of pairwise nucleotide differences between sequences.

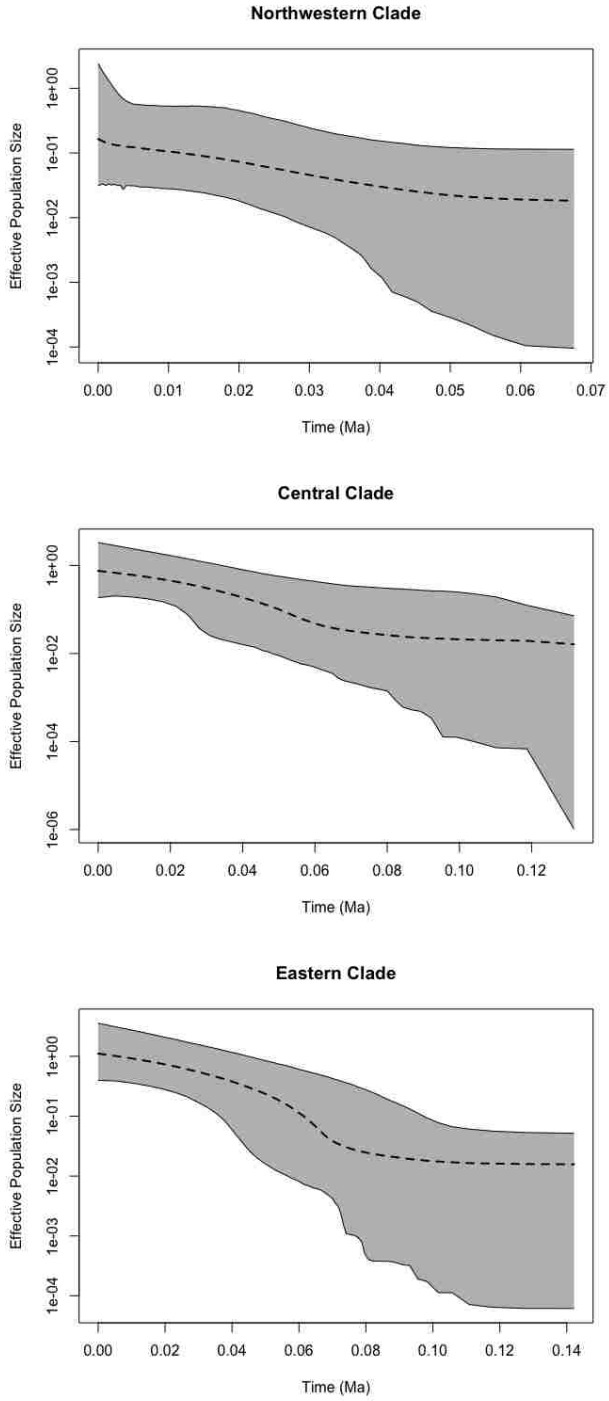


Figure 5. Extended Bayesian Skyline Plots for the Northwestern, Central and Eastern Cytb clades. Analyses are based on sequence data for seven genes (1 mtDNA and 6 nuDNA). The y-axis represents a log-transformed estimation of the effective population size and the x-axis is time measured in millions of years. Dashed line represents the median estimated population size and soled line (gray shading) are the 95% confidence interval.

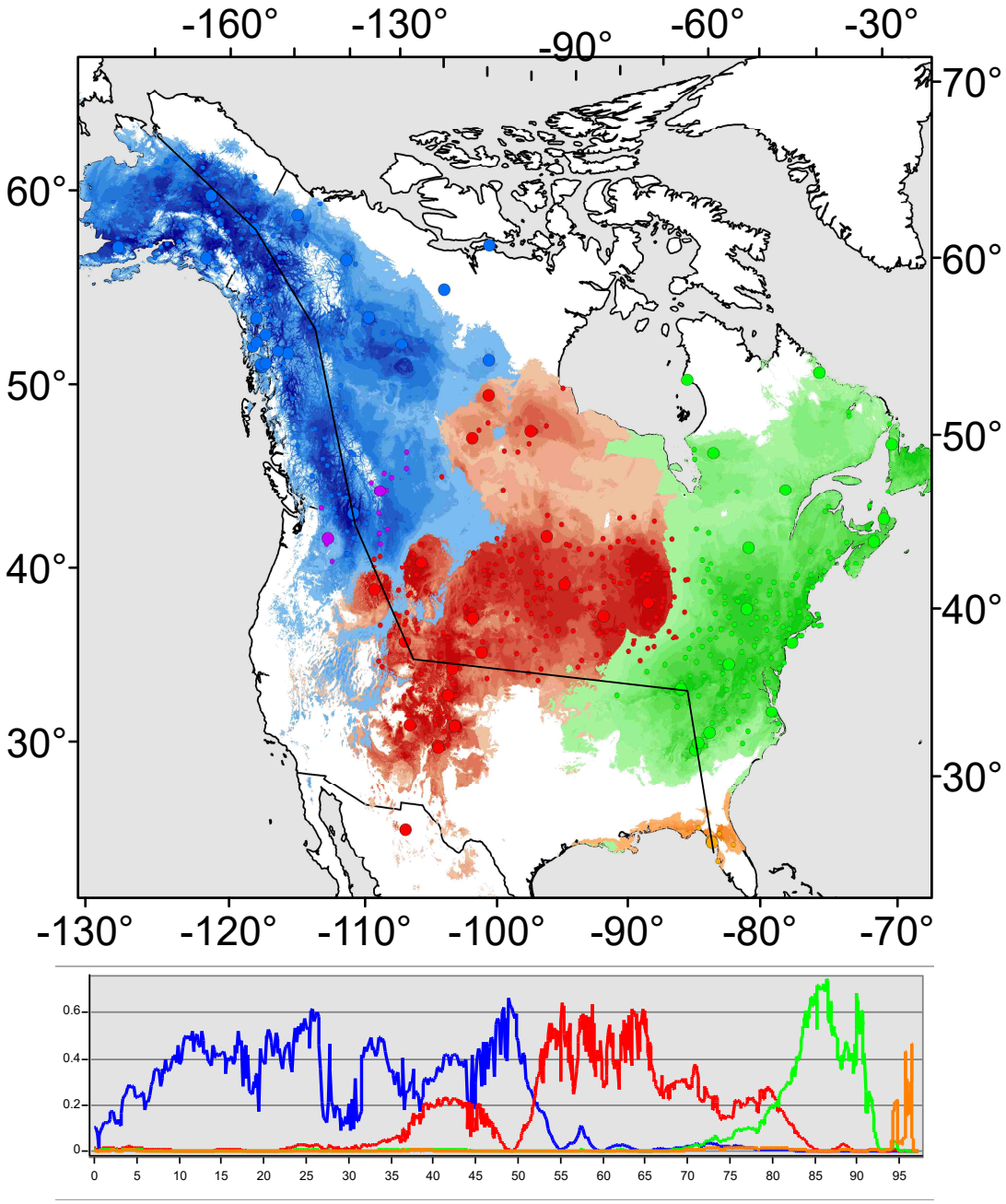


Figure 6. Ecological niche model for *M.* pennsylvanicus present day. Sampled locations are shown by larger circles with colors corresponding to major clades found in Cytb analyses (Figures 1 and 2). Additional smaller dots correspond to known locations where *M. pennsylvanicus* has been reported based on museum specimens, with the colors depicting hypothesized lineage associations.

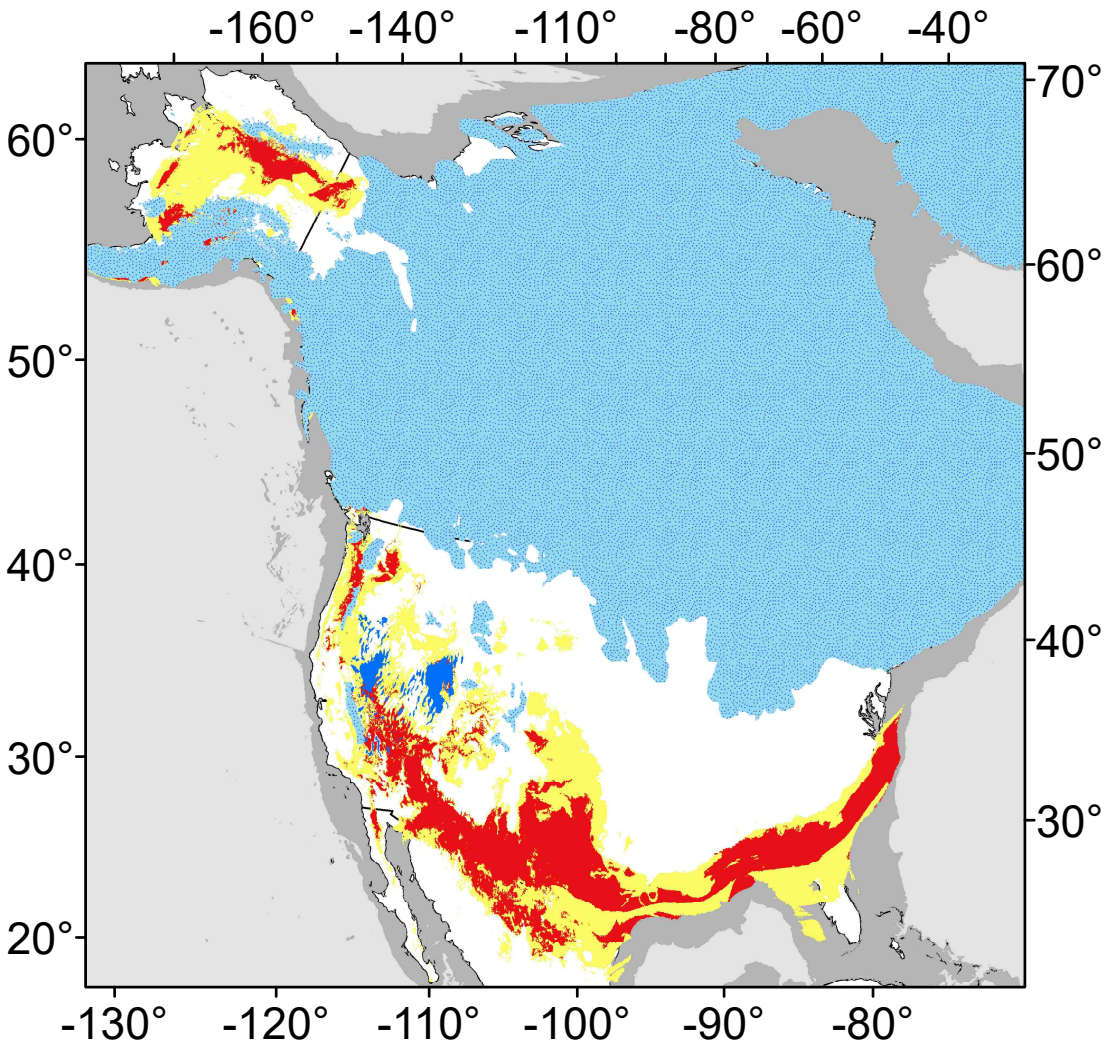
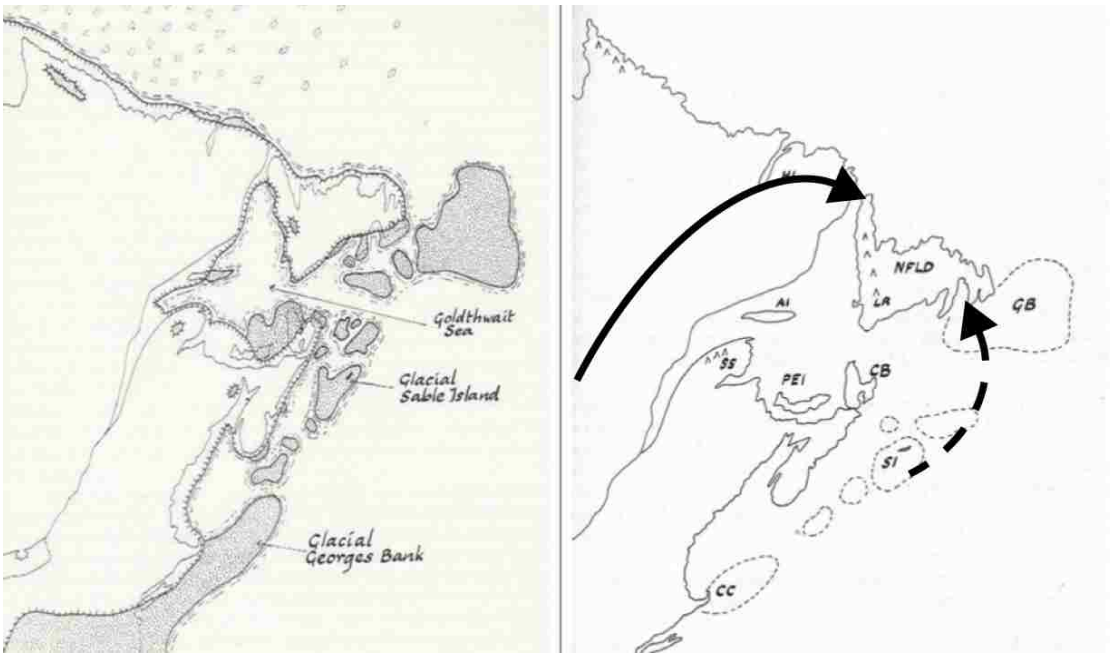


Figure 7. Ecological niche model for *M. pennsylvanicus* during the Last Glacial Maximum. Stippled light blue with blue represent glacial ice coverage and the red and yellow represent suitable habitat for *M. pennsylvanicus* based on higher and lower threshold.



A **B**

Figure 8. Figures A and B are from Pielou (1991) and depict the regions in which coastal refugia are presumed to have been located along the Atlantic Coast of North America during the Wisconsin glaciation (Pielou, 1991). (A). Shown are the boundaries of land, sea and ice during the LGM, shaded areas depict exposed land that may have acted as refugia. (B). Boundaries of modern day coastline with dashed boundaries representing submarine banks that correspond to glacial refugia during the Wisconsin. The two arrows represent the two modes of colonization of Newfoundland, the solid arrow represents the Northern route and dashed arrow the Southern route (Wilkerson, 2010).

Table 1. Primer sequences and annealing temperature. Primers used for amplification and sequencing for *mt*DNA Cytochrome b (Cytb).

Primers	Sequence (5'-3')	Reference	
Cytb (1140 bp)			**51°C**
MSB 05	GACATGAAAAATCATTGTTGTAATTC	Hope *et al*. 2010	
MSB 14	CCCATCTCYGGTTTACAAGAC	Hope *et al*. 2010	
MVZ 04	GCAGCCCCTCAGAATGATATTTGTCCTC	Smith and Patton 1993	
MSB 11	GCCGAGATGTGAACTACGGATG	Hope, A.G. (unpublished)	
MVZ 23	TACTCTTCCTCCACGAAACNGGNTC	Smith and Patton 1993	
MVZ 26	AGATCTTTGATTGTGTAGTAGGGGT	Smith and Patton 1993	
L 649	AAAATTCCCTTCCACCCCTA	Hope, unpublished	
H 885	GCTAGGATTAGGATTGATAGG	Kohli *et al.* 2014	
L 829	GCTTACGCCATCCTACG	Kohli *et al.* 2014	
ETS2 (875 bp)			**63°C**
ETS2 F	AGCTGTGGCAGTTTCTTCTG	Lyons *et al*. 1997	
ETS2 R	CGGCTCAGCTTCTCGTAG	Lyons *et al*. 1997	
FGB (600 bp)			**64°C**
FGBF_M	CGTTTGGATTGGCGGAGTGG	Sawyer and Cook 2016	
FGBR_M	GCACGTACGACAGGGACAACG	Sawyer and Cook 2016	
GHR (460 bp)			**56°C**
GHR-5 F	GGCRTTCATGAYAACTACAAACCTGACYTC	Galewski *et al*. 2006	
GHR-4 R	CTYACYTGRGCATAAAAGTC	Galewski *et al*. 2006	
IRBP (625 bp)			**60°C**
IRBPF P	CCAGGAGGTACTGAGTGAGC	Sawyer, Y.E. 2014	
IRBPR P	GCTGAGTAGTCCATGCTAGC	Sawyer, Y.E. 2014	
LCAT (465 bp)			**54°C**
LCAT F	GTGACAACCAGGGCATCC	Robinson *et al.* 1997	
LCAT R	TCCAGTGTCTTATTGCTGAAG	Robinson *et al.* 1997	
Rag1 (860 bp)			**64°C**
MSB Rag1F_M	GCAGTCTCCTTTAGTTCCAGAC	Sawyer and Cook 2016	
MSB Rag1R_M	CCAACAGGAACAACGTCAAGC	Sawyer and Cook 2016	

Table 2. Molecular diversity indices and population equilibrium tests using 1,111 bp of mtDNA Cytochrome b among all major *M. pennsylvanicus* lineages. Clades are further split based on whether they are located in regions that were covered by ice or maintained suitable habitat during the Pleistocene. Abbreviations include: n = sample size, S = number of segregating sites, H = number of haplotypes, Hd = haplotype diversity, k = average number of nucleotide differences, π = nucleotide diversity, F_s = Fu's Fs, R2 = Ramos-Onsins and Rozas R2, and *D* = Tajima's D. Significance of populations statistics are indicated by asterisks: * $p < 0.05$ and ** $p < 0.0001$ (Fu's Fs $p<0.02$).

Gene/Group	N	S	H	Hd	k	π	*Fs*	R2	*D*
ETS2	28	57	25	0.984	10.873	0.0125	-12.083*	0.0813	-1.075
FGB	28	1	2	0.071	0.071	0.0001	-1.155	0.1856	-1.151
GHR	28	7	7	0.577	0.86	0.0019	-3.43*	0.0758*	-1.568*
IRBP	29	11	12	0.887	2.015	0.0034	-5.611*	0.093	-0.914
LCAT	29	4	5	0.51	0.576	0.0012	-2.138*	0.038	-1.125
Rag1	29	12	10	0.786	1.433	0.0017	-4.933*	0.0674*	-1.745*
Cytb	145	181	98	0.991	24.301	0.0225	-34.333**	0.065	-0.903
Cytb clades									
Northwestern	61	65	38	0.970	5.409	0.0049	-28.245**	0.037**	-2.083*
glaciated	45	60	36	0.986	5.368	0.0048	-33.549**	0.037**	-2.152*
unglaciated	16	13	4	0.750	4.333	0.0039	4.081	0.170	0.409
Central	44	81	32	0.981	9.266	0.0084	-14.806*	0.048*	-1.848*
glaciated	17	35	15	0.985	5.250	0.0047	-8.582*	0.052**	-2.025*
unglaciated	27	54	18	0.966	9.863	0.0090	-3.136	0.080	-1.177
Eastern	38	64	28	0.982	9.243	0.0084	-11.6*	0.062*	1.463*
glaciated	21	42	16	0.971	9.210	0.0084	-3.815	0.091	-0.908
unglaciated	17	31	12	0.949	7.882	0.0071	-1.8725	0.106	-0.573

Table 3. Average pairwise sequence divergence using a Kimura 2-parameter model based on 1,111 bp of Cytb. Average mean divergence between each clade of *Microtus pennsylvanicus* (below diagonal) and net divergence between clades (italics above diagonal). Average within-clade divergence shown along diagonal (in bold).

	Clade/Species	1	2	3	4	*5*
1	Northwestern	**0.006**	*0.01*	*0.034*	*0.036*	*0.042*
2	Central	0.018	**0.009**	*0.03*	*0.036*	*0.037*
3	Eastern	0.042	0.039	**0.009**	*0.033*	*0.012*
4	Florida	0.040	0.041	0.038	**0**	*0.044*
5	*M. breweri*	0.045	0.043	0.018	0.045	**0.001**

Table 4. Results of analyses of molecular variance (AMOVA) for *Microtus pennsylvanicus*. Populations are based on strongly supported Cytb clades (Northwestern, Central, Eastern and Florida). Significance is denoted by * ($p < 0.05$).

	Source of Variation	d.f.	Sum of squares	Variance components	Percentage of variation
Cytb	Among clades	3	1330.028	13.3*	77.63
	Within clades	148	567.277	3.833*	22.37
	Total	151	1897.305	17.1228	

Appendices

Reagents and PCR conditions were: 1 µL of DNA template (variable dependent upon DNA quality), 2.5 µL PCR buffer, 2 µL 25mM $MgCl^2$, 1 µL bovine serum albumin, 0.5µL deoxynucleoside triphosphates, 0.13 µL AmpiTaq DNA Polymerase (Applied Biosystems, Foster City, California), 1 µL of each primer and 8.13 µL of double-distilled H_2O to total 25 µL of reactions. Polymerase chain reactions were performed in a T100 Thermal Cycler (Bio-Rad, USA) that included an initial denaturation at 95°C for 2 minutes followed by 40 cycles of denaturation at 95°C for 15 seconds, annealing for 30 seconds (variable by gene), and extension at 72°C for 45 seconds with a final extension at 72°C for 5 minutes and final cooling at 4°C. All PCR products were visually confirmed by agarose gel electrophoresis and cleaned using either 30% polyethylene glycol precipitation or ExoSAP-IT (Affymatrix, Santa Clara, California) following manufacturer's protocols. Reagents for each 10 µL cycle sequencing reaction were: 2.5 µL of H_2O, 2 µL of BigDye buffer (Applied Biosystems), 1 µL of BigDye v. 3.1 (Applied Biosystems) and 3.5 µL of primer (10 mM) following the conditions of Platt *et al.* (2007). We cleaned cycle sequencing reactions with a 125 mM ethylenediaminetetraacetic acid, sodium acetate and ethanol protocol. Automated sequencing was conducted at the Molecular Biology Facility at the University of New Mexico.

Appendix A. Average pairwise differences using a Kimura 2-parameter model in MEGA 7 for each gene studied among Cytb clades of *Microtus pennsylvanicus*. Standard errors are shown above the diagonal. Mean genetic distance between clades is shown below the diagonal (bold). Mean within-clade distance is shown along the diagonal (in italics). 1) Northwestern; 2) Central; 3) Eastern; 4) Florida.

(a) ETS2

Cytb clade	1	2	3	4
1	*0.008*	0.002	0.003	0.003
2	**0.011**	*0.013*	0.003	0.003
3	**0.015**	**0.016**	*0.013*	0.002
4	**0.015**	**0.015**	**0.002**	-

(b) FGB

Cytb clade	1	2	3	4
1	*0.002*	0.001	0.002	0.002
2	**0.002**	*0.003*	0.001	0.001
3	**0.004**	**0.003**	*0.002*	0.001
4	**0.004**	**0.003**	**0.002**	-

(c) GHR

Cytb clade	1	2	3	4
1	*0.004*	0.002	0.002	0.002
2	**0.005**	*0.006*	0.001	0.002
3	**0.004**	**0.005**	*0.003*	0.002
4	**0.005**	**0.006**	**0.004**	-

(d) IRBP

Cytb clade	1	2	3	4
1	*0.008*	0.002	0.003	0.006
2	**0.007**	*0.005*	0.003	0.006
3	**0.012**	**0.01**	*0.008*	0.005
4	**0.027**	**0.025**	**0.022**	-

(e) LCAT

Cytb clade	1	2	3	4
1	*0.001*	0.001	0.001	0.002
2	**0.002**	*0.001*	0.002	0.003
3	**0.002**	**0.003**	*0.002*	0.002
4	**0.002**	**0.004**	**0.003**	-

(f) Rag1

Cytb clade	1	2	3	4
1	*0.002*	0.001	0.001	0.002
2	**0.002**	*0.001*	0.001	0.002
3	**0.003**	**0.002**	*0.002*	0.002
4	**0.004**	**0.003**	**0.003**	-

Chapter 4

Thesis Conclusion

Climatic fluctuations throughout the Pleistocene played an important role in the speciation and intraspecific diversification of many organisms in North America. My master's thesis focuses on historical processes that catalyzed the proliferation of genetic diversity and shaped the contemporary geographic distribution of the meadow vole (*Microtus pennsylvanicus*). This is the first comprehensive study that utilized range-wide sampling, multilocus techniques, and niche modeling to better understand genetic structure and species limits within this wide spread North American species. My thesis uncovered high genetic variation (comparable to that between sister species within *Microtus*) that was not previously recognized in mtDNA (Plante *et al.* 1987) or morphological studies (Lowry 2002) that were limited in spatial sampling. Additionally, this study revealed multiple incipient species using multilocus analyses and niche modeling and provides vital insight into diversity within this wide ranging species that is critical to the conservation of highly divergent peripheral populations. Moreover, these findings further our understanding of the importance of the climatic fluctuations of the Pleistocene and how these oscillations influenced the diversification and evolution of small mammals throughout North America.

In chapter 2, I performed the first multilocus phylogenetic analyses of *M. pennsylvanicus*, sampling individuals from 20 of 28 subspecies (Hoffmann & Koeppl 1985). Additionally, I explored multiple independently evolving lineages using Bayes Factor species Delimitation. Finally, I explored the implications of modern climate change for peripheral populations along the southern edge of this wide ranging species. I

examined genetic diversity within *M. pennsylvanicus* and compared this variation to morphological subspecies. Phylogenetic analyses strongly supported the presence of 4 distinct clades that are geographically distributed throughout North America. To further characterize genetic diversity within the meadow vole, I constructed a comprehensive phylogeny for 63 species of the genus *Microtus* and evaluated genetic distance between a series of highly supported sister species to place the levels of intraspecific clade divergence I discovered in *M. pennsylvanicus* into the context of interspecific divergence in *Microtus*. Clade distance was within levels of variation between sister species in *Microtus.* With this knowledge, we used BFD analyses to statistically determine if the current taxonomy of this species was reflected in mtDNA and nuDNA analyses. BFD consistently and decisively supported a scenario in which *M. pennsylvanicus* was composed of multiple species. We suggest taxonomy should reflect 3 independent species: 1) a western species that should be *M. drummondii*; 2) an eastern species that will retain the binomial name *M. pennsylvanicus*; 3) and finally the highly restricted populations in coastal Florida should be elevated to *M. dukecampbelli*. Lastly, the nominal species *M. breweri* is minimally distinctive and appears to be conspecific with *M. pennsylvanicus*, however, further analyses using nuclear data should be incorporated to definitively determine the history and status of this insular population.

In chapter 3, I characterize the demographic and evolutionary history of *M. pennsylvanicus*, emphasizing the importance of the dynamic environmental fluctuations of the Pleistocene and the role they played in the contemporary diversification and geographic distribution of the meadow vole. During this vibrant epoch, climatic oscillations resulted in the isolation of meadow voles extended periods of time in

multiple glacial refugia. Previous studies of other species in *Microtus* have shown that populations that persist in separate refugia often undergo rapid diversification that can lead to speciation (Martínková & Dudich 2003; Tougard *et al.* 2008; Krystufek *et al.* 2009). Within this study of *M. pennsylvanicus,* I utilized range wide sampling along with demographic and niche modeling analyses, to better understand the divergence within this species and explore the influence of historical biogeographic events that occurred proceeding the Last Glacial Maximum. Our data support a scenario in which each clade diverged during various glacial and interglacial periods that forced populations to persist in separate refugia. During these periods of isolation, populations rapidly accumulated unique genetic variation. Additionally, we identified two localities that are experiencing secondary contact between the recently diverged Central and Northwestern clades. The eastern and western split between the newly described species *M. drummondii* and *M. pennsylvanicus* coincides with the Last Interglacial, a split that is common in other transcontinental species (Burbrink *et al*. 2008; Puckett *et al.* 2015). Our most distant species, *M. dukecampbelli*, is likely a relict of a more widespread population along the gulf coast. Furthermore, ENM's corroborated fossil evidence of a more widespread distribution south of the current range of *M. pennsylvanicus.* Finally, ENM's of each lineage corroborated genetic data, with each clade occupying a distinct environment and further supports species recommendations in chapter 2.

This study provides vital insight into the response of this species to historic climatic fluctuations, from which we can begin to infer some of the implications of modern climate change. As climate change and conservation continue to gain more attention, it will remain critical that well-informed systematics direct management

strategies for southern peripheral populations to preserve their distinctive genetic variation. The effects of changing climate patterns and human habitat modification have already resulted in the extirpation of relictual populations (List *et al.* 2010) along the southern periphery, serving as a reminder of the vulnerability of *M. dukecampbelli*. Additionally, a clearer understanding of the evolutionary processes that resulted in the distinct clades within the *M. pennsylvanicus* complex enable us to determine where they may have persisted historically and provides a basis for forecasts of how they may respond to warmer temperatures in the future.

References

Burbrink, F.T., Fontanella, F., Alexande,r P.R., Guiher, T.J., & Jimenez, C. (2008) Phylogeography across a continent: The evolutionary and demographic history of the North American racer (Serpentes: Colubridae: Coluber constrictor). *Molecular Phylogenetics and Evolution*, **47**, 274–288.

Hoffmann, R.S., & Koeppl, J.W. (1985) Zoogeography. *Biology of New World Microtus.* (ed. by R.H. Tamarin) pp. 85-115. American Society of Mammalogist. Shippensburg, PA.

Krystufek, B., Buzan, E. V., Vohralik, V., Zareie, R., & Ozkan, B. (2009) Mitochondrial cytochrome b sequence yields new insight into the speciation of social voles in south-west Asia. *Biological Journal of the Linnean Society*, **98**, 121–128.

List, R., Pergams, O.R.W., Pacheco, J., Cruzado, J., & Ceballos, G. (2010) Genetic divergence of *Microtus pennsylvanicus chihuahuensis* and conservation implications of marginal population extinctions Genetic divergence of Microtus pennsylvanicus chihuahuensis and conservation implications of marginal population extinctions. **91**, 1093–1101.

Lowry, J.M. (2002) Systematics of northwestern meadow vole (Microtus pennsylvanicus) subspecies, with emphasis on the island endemic (*M. p. shattucki*, Howe 1901) in Penobscot Bay, Maine. *Thesis, The University of Maine*, .

Martínková, N. & Dudich, A. (2003) The fragmented distribution range of Microtus tatricus and its evolutionary implications. *Folia Zoologica*, **52**, 11–22.

Plante, Y., Boag, P.T., & White, B.N. (1987) Macrogeographic variation in mitochondrial DNA of meadow voles (*Microtus pennsylvanicus*). *Canadian Journal of Zoology*, **67**, 158–167.

Puckett, E.E., Etter, P.D., Johnson, E.A., & Eggert, L.S. (2015) Phylogeographic analyses of American black bears (*Ursus americanus*) suggest four glacial refugia and complex patterns of postglacial admixture. *Molecular Biology and Evolution*, **32**, 2338–2350.

Tougard, C., Brunet-Lecomte, P., Fabre, M., & Montuire, S. (2008) Evolutionary history of two allopatric *Terricola* species (Arvicolinae, Rodentia) from molecular, morphological, and palaeontological data. *Biological Journal of the Linnean Society*, **93**, 309–323.